

2328 + 1415

190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*

117: em_estpl6:*
118: em_estpl7:*
119: em_estpl8:*
120: em_estpl9:*
121: em_estpl10:*
122: em_estpl11:*
123: em_estpl12:*
124: em_estpl13:*
125: em_estpl14:*
126: em_estpl15:*
127: em_estpl16:*
128: em_estpl17:*
129: em_estpl18:*
130: em_estpl19:*
131: em_estpl20:*
132: em_estpl21:*
133: em_estpl22:*
134: em_estpl23:*
135: em_estpl24:*
136: em_estpl25:*
137: em_estpl26:*
138: em_estpl27:*
139: em_estpl28:*
140: em_estpl29:*
141: em_estpl30:*
142: em_estpl31:*
143: em_estpl32:*
144: em_estpl33:*
145: em_estpl34:*
146: em_estpl35:*
147: em_estpl36:*
148: em_estpl37:*
149: em_estpl38:*
150: em_estpl39:*
151: em_estpl40:*
152: em_estpl41:*
153: em_estpl42:*
154: em_estpl43:*
155: em_estpl44:*
156: em_estpl45:*
157: em_estpl46:*
158: em_estpl47:*
159: em_estpl48:*
160: em_estpl49:*
161: em_estpl50:*
162: em_estpl51:*
163: em_estpl52:*
164: em_estpl53:*
165: em_estpl54:*
166: em_estpl55:*
167: em_estpl56:*
168: em_estpl57:*
169: em_estpl58:*
170: em_estpl59:*
171: em_estpl60:*
172: em_estpl61:*
173: em_estpl62:*
174: em_estpl63:*
175: em_estpl64:*
176: em_estpl65:*
177: em_estpl66:*
178: em_estpl67:*
179: em_estpl68:*
180: em_estpl69:*
181: em_estpl70:*
182: em_estpl71:*
183: em_estpl72:*
184: em_estpl73:*
185: em_estpl74:*
186: em_estpl75:*
187: em_estpl76:*
188: em_estpl77:*
189: em_estpl78:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	537	13.1	603	97	AW957049	AW957049 EST369119
2	430.4	10.5	687	106	BE281514	BE281514 601155083
3	428.4	10.4	920	135	BE78623	BE78623 601466167
4	414.8	10.1	420	105	BE245688	BE245688 TCAP1D31
5	334.8	8.1	348	105	BE247025	BE247025 TCAP1D63
6	331.6	8.1	365	105	BE245906	BE245906 TCAP1D34
7	271.4	6.6	284	6	AA367842	AA367842 EST79262
8	271.4	6.6	877	22	AI553743	AI553743 tn28c09.x
9	242.4	5.9	463	93	AW681286	AW681286 EST00004
10	233.4	5.7	344	4	AA223868	AA223868 zr10h12.r
11	219.2	5.3	666	28	AL047744	AL047744 DRP2P586P
12	212.2	5.2	696	106	BE260791	BE260791 601150285
13	210.4	5.1	386	12	AA795800	AA795800 v725h09.r
14	207	5.0	211	6	AA368926	AA368926 EST80222
15	203.8	5.0	636	37	AV611045	AV611045 AV611045
16	203.8	5.0	724	135	BE744312	BE744312 601576526
17	202.2	4.9	660	134	BE047675	BE047675 tz40d09.y
18	200.2	4.9	733	107	BE409700	BE409700 601301556
19	199.8	4.9	407	37	AV596131	AV596131 AV596131
20	198.8	4.8	497	169	AQ001382	AQ001382 HS_5394.A
21	196.8	4.8	461	16	AI092636	AI092636 qa35d08.x
22	188.2	4.6	538	28	AL119940	AL119940 DRP2P761B
23	188.2	4.6	546	134	BE030526	BE030526 1286628.x
24	186.2	4.5	555	23	AI696515	AI696515 vx51f02.x
25	184.4	4.5	591	96	AW872408	AW872408 h191g10.x
26	184.4	4.5	536	136	BE808865	BE808865 214163.MA
27	183	4.5	541	135	BE757540	BE757540 212005.MA
28	182.4	4.4	567	37	AV614826	AV614826 AV614826
29	181.8	4.4	514	38	AV666876	AV666876 AV666876
30	179.8	4.4	603	4	AA266222	AA266222 mz53a06.r
31	178.4	4.3	412	136	BE853746	BE853746 ux20d12.y
32	177	4.3	561	1	AA066758	AA066758 mm09f03.r
33	176.4	4.3	855	19	AI326356	AI326356 mm09f03.y
34	176.4	4.3	855	19	AI326356	BE914826 601667851
35	172.8	4.2	560	137	BE914826	AI325368 m130h09.y
36	172.8	4.2	940	19	AI325368	AI716140 UI-R-Y0-a
37	170.2	4.1	320	24	AI716140	BE706956 QVO-HT036
38	169.6	4.1	366	111	BE706956	AA464876 aa93b10.r
39	167.2	4.1	620	7	AA464876	AA368924 EST80250
40	167	4.1	344	6	AA368924	BE36616 601066875
41	166.2	4.0	624	109	BE536616	BE706950 QVO-HT036
42	164.6	4.0	348	111	BE706950	BE706947 QVO-HT036
43	164.4	4.0	383	111	BE706947	BE226853 us73d11.y
44	163.4	4.0	468	105	BE226853	N64188 yy59e10.r1
45	162.4	4.0	478	143	N64188	

ALIGNMENTS

RESULT 1
AW957049 AW957049 603 bp mRNA EST 01-JUN-2000
LOCUS EST369119 MAGC resequences, MAGD Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION AW957049
VERSION AW957049.1 GI:8146732
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

QY	3701	gccaggattacaactgaggtggtctcttctcccggtgctctgccagagggctgagaccctg	3760
Db	241	gccaggtatttacaacttgggtgtctcttctccgggtgccttggccacaggggctgagaccctg	3820
QY	3761	ggtctctccaagatgaagacatttgaggaaatcccatgaccccaacagacctacaaggc	3820
Db	301	ggttctctccaggatgaagacatttgaggaaatcccatgaccccaacagacctacaaggc	3880
QY	3821	tctgtggaacaccagacagacagtcagtggtgtgctgtgagagagtttgagcagata	418
Db	360	tctgtggaacaccagacagacagtcagtggtgtgctgtgagagagtttgagcagata	3914
QY	3881	gagacaggcatagacaagaagcgccttcagct	452
Db	419	gagacaggcatagacaagaagcgccttcagct	

RESULT 4
 BE245688
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

BE245688 420 bp mRNA
 TCBAP103161 Pediatric pre-B cell acute lymphoblastic leukemia
 Baylor-HGSC project-TCBA Homo sapiens cDNA clone TCBAP3161, mRNA
 sequence.
 BE245688.1 GI:9097435
 EST
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 420)
 Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman Jr., F.R., Muzny, D.,
 Bouck, J., Gibbs, R.A. and Margolin, J.F.
 Pediatric Leukemia cDNA Sequencing Project
 Unpublished (2000)
 Contact: Dr. Judith F. Margolin
 Human Genome Sequencing Center at Baylor College of Medicine and
 Texas Children's Cancer Center
 One Baylor Plaza, Houston, TX 77030, USA
 Tel: 713 770 4536
 Fax: 713 770 4038
 Email: jmargin@tccc.org
 Seq primer: M13 primer.
 Seq primer: M13 primer.

```

1. Location
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCBAP3161"
/clone_lib="pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project-TCBA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/notice="Vector: lambda PSB; Site_1: BamHI; Site_2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGACTCTGAGCGCGCCGACAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second
strand was primed with a BamHI-dC primer
[5'AGAGACTCTGGATCCGCGCGCCGAATAATAATAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and
XhoI and directionally cloned into the BamHI and XhoI
sites of lambda PSB vector. Library was constructed by Wel
round of normalization. Library was constructed by Wel
YU"

```

BASE COUNT	88 a	122 c	139 g	0.7 c
ORIGIN				
Query Match		10.1%	Score 414.8;	DB 105; Length 420;
Similarity		99.0%	Pred. No. 1.1e-81;	

Query Match

[illegible]

RESULT	3	EST	20-SEP-2000
BE778623		920 bp	mrna
LOCUS		NIH_MGC_67	Homo sapiens cDNA Clone IMAGE:3869200 5',
DEFINITION		601466167E1	
		mrna sequence.	

BE778623
 BE778623.1 GI:10199821
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NTH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
 1 (bases 1 to 920)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, ph.D.
 Tel: (301) 496-1550
 Email: Robert_Strausberg@nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM605 row: f column: 17
 High quality sequence stop: 716.
 High quality sequence qualifiers

```

Location/Query:
1. .920
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3869200"
/clone_lib="NIH_MGC_67"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

BASE COUNT      236 a      240 c      268 g      176 t
ORIGIN

Query Match      10.4%; Score 428.4; DB 135; Length 920;
Best Local Similarity 99.3%; Pred. No. 1.2e-84;
Matches 451; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 3461 atgtgaactgctggtccggagagaccccaaggagacctgcattctcggagctgtaggag 3520
1 ATGCTGAACCTGCTGGTCCGGAGACCCNCCAGCGAGACCTGTCATCTCGGAGCTGTGTGGAG 60

Db 1 atctgggggaacctgctccaggcaggcggcctgcagaaggaagagggtctgcattgccc 3580
3521 atctgggggaacctgctccaggcaggcggcctgcagaaggaagagggtctgcattgccc 120
61 ATCTGGGGGAGACCTGCTCCAGGCGAGGGGCTGCAAGAGGAAGAGGAGGTCTGCATGGCC 180
3581 ccgcgcagctctcgagctcagaagaggcagctctctcgagggtgtccaccattggcccta 3640
QY 121 CCGCGCAGCTCTCAGAGCTCAGAAGAGGGCGAGCTTCTCCAGGCTGTCCACCATGGCCCTA 180
Db 121 CCGCGCAGCTCTCAGAGCTCAGAAGAGGGCGAGCTTCTCCAGGCTGTCCACCATGGCCCTA 180
3641 cacatcgccaggctcagctgaggacagcccgcaagcctcgacgcgcacagcctggccc 3700
QY 181 CACATCGCCCNAGCTCAGGCTGAGGACAGCCCGCCAAAGCCTTCAGCGCACAGCCTGGCC 240

```


Moreno-Palanco, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Kyder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Ding, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Dinkar, D., Feng, D.-F., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Hu, J.S., Greene, J.M., Gruber, J., Weissner, P.S., Olsen, H., Raymond, L., Kunsch, C., Hungjun, Xu, C., Yu, G.-L., Ruben, S.M., Dillon, P.J., Fannon, Wei, Y.F., Wang, J., Haseltine, W.A., Fields, C., Fraser, C.M. and Jenter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)

Other ESIS: THC85507
96026280

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
7912 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
seq primer: M13-21

FEATURES
source

organism="Homo sapiens"
 db_xref="ATCC (inhost):172752"
 db_xref="taxon:9606"
 db_xref="placenta I"
 clone_lib="placenta I"
 tissue_type="placenta"
 dev_stage="fetus"
 note="Organ: placenta"
 ECoRI: Site2: 91 g
 61 C

BASE COUNT
ORIGIN

|||||TTTCATCGGCAACCACTCAACGTGGTCAACGCT

101
284

2781 9CGGCTTCATGGTGATCGTGGAGTTCTGCAAGTACGGCAAGCTG

EST

AI553743 877 bp Homo sapiens CDNA cDNA

tn28c09.x1 NC12C093 TYROSINE-PROTEIN
similar to gb:X68203 contains element MER22 repet

PRECUSOR (HUMAN), CON-
sequence.

SSION
AI553743

1

1

VERSION	A1553743.1	GI:4486106
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	NCI/NINDS-CCAP http://www.ncbi.nlm.nih.gov/ncicgap. 1 (bases 1 to 877)	
JOURNAL	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project	
COMMENT	(CGAP/BTCAP), Tumor Gene Index Unpublished (1998) Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert.Strausberg@nih.gov Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D. CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D. CDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CCAP clone distribution Information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/bbrp/image/image.html Insert Length: 1499 Std Error: 0.00 Seq Primer: -40UP from Gibco High quality sequence stop: 400 POLYA=No. Location/Qualifiers 1. 877 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:216894" /clone_lib="NCI-CCAP_Brn25" /tissue_type="anaplastic oligodendroglioma" /lab_host="DH10B" /note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'- TGTTACCAATCTGAAGTGGGAGCGGCAGTAGCTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo." 152 a 257 c 260 g . 202 t 6 others	
FEATURES	source	
BASE COUNT		
ORIGIN		

Query Match		6.6%; Score 271.4; DB 22; Length 877;		352; Conservative 0; Mismatches 81; Indels 5; Gaps 3;	
Best Local Similarity 80.4%; Pred. No. 7.7e-50;					
Matches 352; Conservative 0; Mismatches 81; Indels 5; Gaps 3;					
QY	3468	actgctggtccgagagaccccccaaggcagacctgcattctcggagctgtgagatcctgg	3527		
Db	847	AATGTTGGTCCGAAGACCCCAAGCG---GAACCTGATTTTCGGAATGGGGGAAGATCCGG	791		
QY	3528	ggagacctgtccaggcagggccctgcgaagagagagaggtctgcagtcgcccgcgca	3587		
Db	790	GGACTTGTCT-CAGGCCAGGGCCTGCAAGAGG-AGAGGAGGCTTGCAATGCCCCCGCGCA	733		
QY	3588	gctctcagagctcagaagaaggcagcttctcgcaggtgtccaccatggccctacacatcg	3647		
Db	732	GCTCTCAGAGCTCAGAAGAGGGCAGCTTCTCGCAGGTCTCNACCATGGCCCTACACATCG	673		
QY	3648	cccaggtgaagctgagacagcccgcccaagcctgcacagccacagcctggccgcaggt	3707		
Db	672	CCGAGGCTGACGNTGAGGACAGCCGCCNAAGCCCTGCAAGCCACAGGTGCCCGCAGGGT	613		
QY	3708	attacaactgggtgtcctttcccggtgtcctggccagaggggtgagaccgtgggttctt	3767		
Db	612	ATTAACTGGGTGTCTTTTCCCGGGTGCCTTGGCCACAGAGGCTGAGACCGGTTCCT	553		

3317	tttggggtctctctgagagatctctctctgggggccctcccctaccctggggtgcag	3376	
QY			
Db	61	ttttggggtctctctgagagatctctctctgggggccctcccctaccctggggtgcag	3436
QY	3377	atcaatgagagagttctgccagcgctgagagagcacaagatgagggcccccgagctcg	3495
Db	121	ATCAATGAGGAGTCTTGC CAGC - GNTGAGACCGCCACAGGATGATGGGCCCGCGAGCTN	179
QY	3437	ggcactcccgccatcacgcgcgcatcatgtcgaac-tgctgtgtccggagaacccccaaaggcgag	3495
Db	180	GCCACTCTCGGCATACGCCGCATCATGCTGAACNTGCTGGTCCGGAGAGACCCACCAAGCGCAG	239
QY	3496	acctgcatctcg---gagctggtggagagatcctgggagacc-tgctccagggggcagggg	3549
Db	240	ACCTGCACTCTCGGGAGCCCTGGGTGGAGATCCTGTGGGACCTTGCTCAGGACCAAGGG	297
QY			
Db			

RESULT 11	AL047744	666 bp	mRNA	EST	01-MAR-2000			
LOCUS	DRZF586P1721_r1	586	(synonym: hutel)	Homo sapiens	CDNA clone			
DEFINITION	DRZF586P1721, mRNA sequence.							
ACCESSION	AL047744							
VERSION	AL047744.1	GI:4728740						
KEYWORDS	EST.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.							
	1 (bases 1 to 666)							
	Ottewaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and							
	Wiemann,S							
	EST (Ottewaelder, et al.)							
TITLE	EST (Ottewaelder, 1999)							
JOURNAL	Unpublished							
COMMENT	Contact: Ottewaelder B							

HIPS
 Ann Klopferspitz 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 s. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de
 (Martinsried/Germany) within the cDNA
 sequencing consortium of the German Genome Project. No si sequence
 available.
 This clone (DKFZp386p1721) is available at the RZPD in Berlin.
 RZPD; Ressourcenzentrum, Heubnerweg 6, 14059
 clone@rzd.de

```

FEATURES
source
Location/Qualifiers
1..666
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZ5586p171"
/tissue="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/site="Site_1: Not1; Site_2: SalI/MluI"

```

```

BASE COUNT      141 a  197 c  180 g  115 t
ORIGIN
Query Match      5.3%; Score 219.2; DB 28; Length 666;
Best Local Similarity 66.5%; Pred. No. 2.7e-38;
0; Mismatches 158; Indels 0; Caps 0;
Matches 314; Conservative
ccccccagggatgattctt 3102

```

213	Db	GTCAGAAAAGGAGATGCCCGACTCCCTTTGAAGTGGATGGCCCGGMAACCATTTTTCAC	272
213	Db	GTCAGAAAAGGAGATGCCCGACTCCCTTTGAAGTGGATGGCCCGGMAACCATTTTTCAC	272
3284	QY	aagtgctacaccacgcagagtgacgtgctcttgggtgctctctctggagatcttc	3343
3284	QY	aagtgctacaccacgcagagtgacgtgctcttgggtgctctctctggagatcttc	3343
273	Db	AGATATACACAATTACAGACGATGTGGTCTTTCGGTGTCTGTGGGAAATATT	332
273	Db	AGATATACACAATTACAGACGATGTGGTCTTTCGGTGTCTGTGGGAAATATT	332
3344	QY	tctctggggggctcccgtaacctggggtgcagatacaatgagaggtcttcgccagcgctg	3403
3344	QY	tctctggggggctcccgtaacctggggtgcagatacaatgagaggtcttcgccagcgctg	3403
333	Db	TCTTTAGGTGGCTCCCATACCTTGGGGTCAAGATTGATGAAGAAATTTTCTAGGAGATTG	352
333	Db	TCTTTAGGTGGCTCCCATACCTTGGGGTCAAGATTGATGAAGAAATTTTCTAGGAGATTG	352
3404	QY	agagcgcgcacaagagatgagggcccccggagctggtccactcccgcatacgcc	3455
3404	QY	agagcgcgcacaagagatgagggcccccggagctggtccactcccgcatacgcc	3455
393	Db	AAAGAGGAACTAGAAATCGGGCTCTGACTACACTACCCCGAAGATGTACC	444
393	Db	AAAGAGGAACTAGAAATCGGGCTCTGACTACACTACCCCGAAGATGTACC	444

RESULT 10
 AA223868 344 bp mRNA
 AA223868 344 bp mRNA
 EST 11-MAR-1998
 cDNA clone IMAGE:651143 5' similar to gp:Y68203 TYROSINE-PROTEIN
 KINASE RECEPTOR FL14 PRECURSOR (HUMAN); mRNA sequence.
 AA223868
 AA223868.1 GI:1844453
 EST
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 344)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Kriznan, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin
 J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL : contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 1275 Std Error: 0.00
 Seq primer: -28m13 rev1 ET from Amersham
 238.

```

Location/quantity
1. .344
/organism="Homo sapiens"
/organism="GDB:588193"
/db_xref="taxon:9606"
/db_xref="IMAGE:651143"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/clone_lib="neuroepithelial cells"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial"
/lab_host="SOLR (kanamycin resistant)"
/notes="Organ: brain; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Uninduced, exponentially growing neuroepithelial cells (Ntera-2/cl.D1). Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5, adaptor sequence: 5' GAATTCGCACGAG 3' -3, adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3' 66 r

```

```

65 a 92 c 115 g 00 c
SE COUNT
GIN
5.7% Score 233.4; DB 4; Length 344;
Query Match 95.0%; Pred. No. 1.7e-41; Indels 6; Gaps 4;
Best Local Similarity
Matches 283; Conservative 0; Mismatches 9;
3257 tggatggccctgaaagcattctgcacaaggtgtacaccacgacagatgacgtgtgctcc 3316
3257 tggatggccctgaaagcattctgcacaaggtgtacaccacgacagatgacgtgtgctcc 60

```

3257 tggatggccctgaaagc

Query Match 5.2%; Score 212.2; DB 106; Length 696;

FEATURES	
source	1..386
	/organism="Mus musculus"
	seq primer: -28ml3 rev1 ET from Amersham
	High quality sequence stop: 15.
	Location/Qualifiers

us-09-375-248-1.lst

Tue Jan 23 12:06:49 2001

```

/strain="NIH/Swiss"
/db_xref="taxon:10090"
/clone_image="1223489"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: heart; Vector: pBluescript SK-; Site: 1;
EcORI; Site: 2; XhoI; Cloned unidirectionally. Primer:
Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts.
Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5'
adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTGTGTTTTTTTTT 3'."
BASE COUNT      112 a 103 c 97 g 74 t
ORIGIN
Query Match      5.1%; Score 210.4; DB 12; Length 386;
Best Local Similarity 77.9%; Pred. No. 2.2e-36;
Matches 304; Conservative 0; Mismatches 81; Indels 5; Gaps 4;
QY 3583 gcgcagctctcagagctcagaagagcagctctcgcaggtgtccaccatgcccctaca 3642
Db 1 GCACAGCTCTCAGAGCTCCGAGGAGGATGGCTTCATCGAGGATCCA-CACAGCTCTACA 59
QY 3643 catcgccagctgagctgagagacagccgcaagctcagcgccacagcctggccgc 3702
Db 60 TATCAACGAGCAGCAGCTGATGATGCTCTCACCAGCATGCTTGGCCACAGCCTGGCAGC 119
QY 3703 caggtattacaactgggtgtcctctcccggtgctgagcagaggggtgagaccgtgg 3762
Db 178 CTCTTCCAGGATGAGACATTTCTAGAAATGGCCATGACCCCTACAACTCAAGCCTC 237
QY 3763 ttctccaggaataagacatttggagaaattcccccattgaccccaagcaccataaaggctc 3822
Db 178 CTCTTCCAGGATGAGACATTTCTAGAAATGGCCATGACCCCTACAACTCAAGCCTC 237
QY 3823 tctggaacacagacagcagtgaggtgctgctgagcagagaggttg-agcaagatag 3881
Db 238 CATGNTAAACAGACAGACAGCGGATGTGCTGGCTCAGAAAGAGTTTGAAGAGCTAG 297
QY 3882 agagcagcagatagacaagaagcggcttcagctgtaagaagcctggcagaatgtgactg 3941
Db 298 AAGCAGGCGATAGACAGACAGGAGCAGCTTCA-CTGTAAAGTTCTGGCCAGCAGCATGGATA 356
QY 3942 tgaccagggcacacccctgactcccaaggga 3971
Db 357 TTCCAGAGGGAACCTGAACCCCAAGGGA 386
RESULT 14
AA368926 211 bp mRNA EST 21-APR-1997
LOCUS ESP80252 Placenta I Homo sapiens cDNA similar to tyrosine kinase,
DEFINITION receptor FL74, class III, mRNA sequence.
ACCESSION AA368926
VERSION AA368926.1 GI:2021244
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
O., Sutton,G., Blake,J.A., Brandon,R.C., Man,Wai,C., Clayton,R.A.,
Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,

```

Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A., He,W.M.,
Hu,J.S., Greene,J.M., Li,H., Weisner,P.S., Olsen,H., Raymond,L.,
Kunsch,C., Hung,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
Other ESTs: THC85507
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerl@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13-21.

Location/Qualifiers

1..211
/organism="Homo sapiens"
/db_xref="AFCC (Inhost):173626"
/db_xref="taxon:9606"
/clone_lib="placenta I"
/tissue_type="placenta"
/dev_stage="fetus"
/note="Organ: placenta; Vector: pBluescript SK-; Site: 1;
EcORI; Site: 2; EcORI"
46 a 71 c 36 t 4 others

BASE COUNT
ORIGIN

Query Match 5.0%; Score 207; DB 6; Length 211;
Best Local Similarity 98.1%; Pred. No. 1.1e-35;
Matches 207; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1623 tggctcccaagtgaggccaggtgagcgctcatcttcttctgtgaccaccatcc 1682
Db 1 TGTGTCCTCAAGGTGGCCAGGATGAGCGGCTCATCTACTTCTATGTGACCACTCC 60
QY 1683 ccgacggctccaccatcgaaatccagccatccgagagagctactaagggccagcggtgc 1742
Db 61 CCGACGGCTTCACCATCGANTCCAGCCATCCGAGGAGCTACTAGAGGCCAGCGGTGC 120
QY 1743 tctgagctgccaagccagcagctacaagtcagcagcatctgcgtgtgacgcctcaacc 1802
Db 121 TCTGAGCTGCCAAGCCGACAGCTACAGTACGAGCATCTNCGCTGTGTACCGNCTCAACC 180

QY 1803 tgtccagctgcagcagtgatgcgacgggaacc 1833
Db 181 TGTCACGCTGCACGATGCGCACGGGAACCC 211

RESULT 15
AV611045 636 bp mRNA EST 30-AUG-2000
LOCUS AV611045 Bos taurus lung fetus Bos taurus cDNA clone E110036C04 5',
DEFINITION mRNA sequence.
ACCESSION AV611045.1 GI:9746715
VERSION EST.
KEYWORDS Bos.taurus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae.
REFERENCE 1 (bases 1 to 636)

AUTHORS Sugimoto, Y., Hirotsune, S., Takasuga, A., Itoh, R., Jitohzono, A. and Suzuki, H.

TITLE bovine cDNA sequencing
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@ocn.ocn.ne.jp

Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES
source
1. .636
Location/Qualifiers
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="E11U036C04"
/clone_lib="Bos taurus lung fetus"
/tissue_type="lung"
/dev_stage="fetus"
/lab_host="DH10B"
/note="Vector: pZ11; Site_1: SalI; Site_2: NotI; Poly A was deleted from a NotI site"
BASE COUNT 138 a 203 c 177 g 118 t
ORIGIN

Query Match 5.0%; Score 203.8; DB 37; Length 636;
Best Local Similarity 65.1%; Pred. No. 7.le-35;
Matches 317; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 3041 ccgctgaccatgaagatctgtctgtacagcttccaggtgcccagaggtgaggttc 3100
Db 87 CAGCTCTCTCCCAAGACCTGCTGCTCGGCTACAGGTGCCCCGAGGCATGGAGTAT 146
QY 3101 ctggtctcccgaaagtgcacacagagacctggctgcggaaacattctgtctcggaa 3160
Db 147 CTTGCCCTCCCAAGAGTGCATCCACCGGACCTGGCCGCCAGGAACGCTCCTGTGACGGAG 206
QY 3161 agcagctgtgaagatctgtgacttggcttgccttgcggagacatctacaagacctgac 3220
Db 207 GACAACTGATGAGATCGCGGACTTCGGTCTTGTCTCGAGACATCCACCACATCGACTAC 266
QY 3221 tacgtccgcaaggcgagtgcccggctgcccctgaagtgcagtgatggcccctgaagcatcttc 3280
Db 267 TATAAAGACAAACCAACGGCCGACTGCCCGTCAATGGATGGACCGAGGCTTGT 326
QY 3281 gacaaggtgtacaccagcagagtgacgtgtgtctcttgggtgcttctcttggagatc 3340
Db 327 GACCGGATCTACACCCACAGAGCGAGTGTGTTTGGGGTGTCTCTCTGGGAAATC 386
QY 3341 ttctctgggggctcccccgtaccctggggtgcagatcaatgagaggttctgcccagcg 3400
Db 387 TTCACCTGGGCGGCTCCCCATACCTGGGGTCC---CCGTGGAGGAGCTTTTCAAGCTG 443
QY 3401 ctgagagcggcacaagatgagggcccgaggctggccactcccccatacgcgcacac 3460
Db 444 CTGAAGGAGGGTTCATCGTATGGACAGCCAGTAACTGCACCAACGAGCTTCTACATGATG 503
QY 3461 atgctgaactgctggtccgagaccccccaagcagagacctgcatctcggagctggtggag 3520
Db 504 ATGAGAGATTGCTGGCAGCGGGTCCCTCTCACAGACCCACCTTCAAGCAGCTGGTGGAA 563
QY 3521 atcctgg 3527
Db 564 GACCTGG 570

Search completed: January 23, 2001, 03:35:38
Job time: 8492 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2001, 01:43:56 ; Search time 102.84 Seconds
(without alignments)
6442.338 Million cell updates/sec

Title: US-09-375-248-1
Perfect score: 4111

Sequence: 1 ccagcgcagcgccgagaga.....tcacagacacagctactaa(4111)

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2.6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2.6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2.6/ptodata/1/ina/6_COMB.seq.*
- 4: /cgn2.6/ptodata/1/ina/PCUS_COMB.seq.*
- 5: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match	Length	ID	Description
1	4111	100.0	4795	1	US-08-340-011-3	Sequence 3, Appli
2	4111	100.0	4795	3	US-08-901-710-3	Sequence 3, Appli
3	3912.4	95.2	4195	1	US-08-340-011-1	Sequence 1, Appli
4	3912.4	95.2	4195	3	US-08-901-710-1	Sequence 1, Appli
5	3912.4	95.2	4416	3	US-08-795-430-1	Sequence 1, Appli
6	3912.4	95.2	4425	1	US-08-222-616-31	Sequence 31, Appl
7	3912.4	95.2	4425	4	PCT-US95-04228-31	Sequence 31, Appl
8	3912.4	95.2	9108	4	PCT-US95-04228-45	Sequence 45, Appl
9	1040.2	25.3	6827	1	US-08-222-616-17	Sequence 17, Appl
10	1040.2	25.3	6827	4	PCT-US95-04228-17	Sequence 17, Appl
11	716.2	17.4	5470	2	US-08-443-861-1	Sequence 1, Appli
12	713	17.3	5406	1	US-07-813-593-3	Sequence 3, Appli
13	713	17.3	5406	1	US-07-977-451-5	Sequence 5, Appli
14	713	17.3	5406	1	US-07-946-507-3	Sequence 3, Appli
15	713	17.3	5406	1	US-08-252-517-7	Sequence 5, Appli
16	713	17.3	5406	1	US-07-906-397A-5	Sequence 5, Appli
17	713	17.3	5406	1	US-08-601-891-5	Sequence 5, Appli
18	713	17.3	5406	2	US-09-021-324-5	Sequence 5, Appli
19	713	17.3	5406	4	PCT-US92-02750-7	Sequence 7, Appli
20	713	17.3	5406	4	PCT-US92-05401-5	Sequence 5, Appli
21	713	17.3	5406	4	PCT-US92-09893-5	Sequence 5, Appli
22	681.4	16.6	4236	1	US-08-810-116-7	Sequence 7, Appli
23	681.4	16.6	4236	2	US-07-930-548A-7	Sequence 7, Appli
24	242.8	5.9	3453	1	US-07-813-593-1	Sequence 1, Appli
25	242.8	5.9	3453	1	US-07-977-451-1	Sequence 1, Appli
26	242.8	5.9	3453	1	US-07-946-507-1	Sequence 1, Appli
27	242.8	5.9	3453	1	US-08-252-517-1	Sequence 1, Appli
28	242.8	5.9	3453	1	US-07-906-397A-1	Sequence 1, Appli

29	242.8	5.9	3453	1	US-08-601-891-1	Sequence 1, Appli
30	242.8	5.9	3453	2	US-09-021-324-1	Sequence 1, Appli
31	242.8	5.9	3453	4	PCT-US92-02750-1	Sequence 1, Appli
32	242.8	5.9	3453	4	PCT-US92-05401-1	Sequence 1, Appli
33	242.8	5.9	3453	4	PCT-US92-09893-1	Sequence 1, Appli
34	242.8	5.9	3521	1	US-08-222-299-1	Sequence 11, Appli
35	242.8	5.9	3521	2	US-08-434-878-1	Sequence 1, Appli
36	242.8	5.9	3521	4	PCT-US95-03718-1	Sequence 1, Appli
37	228.4	5.6	4465	1	US-08-180-195-1	Sequence 1, Appli
38	228.4	5.6	4465	1	US-08-477-329-1	Sequence 1, Appli
39	228.4	5.6	4465	2	US-08-475-458-1	Sequence 1, Appli
40	228.4	5.6	4465	3	US-08-980-400-1	Sequence 1, Appli
41	226.8	5.5	5427	1	US-08-168-917-1	Sequence 1, Appli
42	226.8	5.5	5427	2	US-08-460-510-1	Sequence 1, Appli
43	226.8	5.5	5427	2	US-08-460-490-1	Sequence 1, Appli
44	226.8	5.5	5427	3	US-08-462-728-3	Sequence 3, Appli
45	226.8	5.5	5427	4	PCT-US92-00730-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-340-011-3
; Sequence 3, Application US/08340011
; Patent No. 5776755
; GENERAL INFORMATION:
; APPLICANT: Alltalo, et al.
; TITLE OF INVENTION: FLT4, A NOVEL RECEPTOR TYROSIN KINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340.011
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/959,951
; FILING DATE: 09-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: 36,989
; REFERENCE/DOCKET NUMBER: 32267
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..4111
US-08-340-011-3

Query Match 100.0%; Score 4111; DB 1: Length 4795;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ccacgcagcggccgagatgagcggggcgccgctgctcgactgtggctctg 60
Db 1 CCACGGCAGCGCGGAGATGACGGGGCGCGCTGTGCTGGACTGTGGCTCTG 60
Qy 61 cctggactcctgagcggcctggtgagtgactactccatgacccccccgacctgaacat 120
Db 61 CCTGGACTCTCGACGGCTGTGTGAGTGGCTACTCCATGACCCCGGACCTTGAACAT 120
Qy 121 cacggaagatcacagtcacgaccggtagacacctgacagccttccatctcctgagggaca 180
Db 121 CACGGAGAGTCAACATCATGACACCGGTGACAGCTGTCCATCTCTGACGGGACA 180
Qy 181 qcacccctcagtggtctggtccagagctcagggagcggccagccagcagagacaagga 240
Db 181 GCACCCCTCGAGTGGCTTGGCCAGGAGCTCAGGAGCGCCAGCCACCGAGACAAGGA 240
Qy 241 cagcgagacacggggtggtgcgagactgcgaggggcagagcgccagcctactgcaa 300
Db 241 CAGCGAGACACGGGGTGTGTGAGAGCTGCGAGGGACACAGACGCCCTTACTGCAA 300
Qy 301 ggtgttgcctgcacaggtacatgccaacgacacagcgagctacgtctgactactaaa 360
Db 301 GGTGTGTGCTGTGCACAGGTACATGCCAACAGCTGACACAGGAGGAGCTTGTGCTACTACAA 360
Qy 361 gtacatcaagcagcgtcaggggacacggccgcccagctcctacgtgttctgtgagaga 420
Db 361 GTACATCAAGCAGCATCGAGGGCACACGGCCGCGAGCTCCTACGCTGTGTGTGAGAGA 420
Qy 421 ctttgagcagcattcatcaagaagctgacagctcctgttgaacaggaagcagcat 480
Db 421 CTTTGAGCAGCATTCATCAACAGCTGACACGCTCTTGTGTCAACAGGAAGGACGCCAT 480
Qy 481 gtgggtgcccctgtctgtgtccatccccggcctcaatgtcacgctgcgctcgcaaaagctc 540
Db 481 GTGGGTGCCCTGTGTGTGTCATCCCCGGCTCAATGTCTACAGCTGCGCTCGCAAAAGCTC 540
Qy 541 ggtgctgtgcccagcagggaggggtgtgtggatgacccggcggggcatgctcgtgtc 600
Db 541 GGTGTGTGGCCAGACGGGAGGAGGTGTGTGGATGACCGCGGGGCGATGCTCGGTGTC 600
Qy 601 cagcgcactgctcacagatgcccctgactcagtgagagaccacccctgggagaccagga 660
Db 601 CAGCGCACTGTGACAGATGCCCTGTACTTGCAGTGGAGACCACTTGGGAGACCAAGGA 660
Qy 661 cttcctttccaaaccccttctgtgtgcacatcacagggcaacgagctctatgacatccagct 720
Db 661 CTTCTTTTCCAAACCCCTTCTGTGTGCACATCACAGGCAACGAGCTCTATGACATCCAGCT 720
Qy 721 gttgcccaggaagtgcgtggagctgctgttagggagaagctgctcctgaactcacagt 780
Db 721 GTTGGCCAGGAAGTGCCTGGAGCTGCTGCTGAGGAGAGGCTGTGCTTGAACCTGACCCGT 780
Qy 781 gtgggtcaggtttaactcaggtgtcacctttgactggagactaccagggagcagggcaga 840
Db 781 GTGGCTGAGTTTAATCAGGTGTCACTTGTACTGGGACTACCCAGGGAAGCAGGCAGA 840
Qy 841 gcggggtgaagtgggtgcccagcagcgtcccagcagaccacacagaactctccagcat 900
Db 841 CGGGGTGAAGTGGGTGGCCGAGCAGCTCCGACAGCAACCCACAGAACTCTCCAGCAT 900
Qy 901 cctgaccatcacacagtcagccagcagcactgggtcgtatgtgtgcaaggcaacaa 960
Db 901 CCTGACATCCACAGTACGCCAGCACGACTGGGCTGTGTGTGTGCAAGGCCAACAA 960
Qy 961 cggcatccagcagatttcgggagagcacccaggttcaattgtgatgaaaaatcccttcacag 1020
Db 961 CGCATCCACGATTTTCGGGAGAGCACCCAGGTCATTTGTGATGAAATCCCTTTCATCAG 1020
Qy 1021 cgtcagatggtctcaagagaccatcctgagggccagcagagagcagcctggtgaagct 1080
Db 1021 CGTGAGTGGCTCAAGGAGACCCATCTCTGGAGGCCACCGCAGGAGACGAGCTGTGTAAGCT 1080
```


QY 1441 ccggcggcgcagcagcaaacctcatgcccacagtgcccgtagctggagggcggtgaccac 1500
|||||
DB 1441 CCGCGCGCGCAGCAGCAAGACCTCTATGCCACAGTGCCTGTACTGAGGCGCGTGACCAC 1500
|||||
QY 1501 qcaggaTgcgtgaaccccatcagagccttgagacactagaccgaatttgagagga 1560
|||||
DB 1501 GCAGAGTGCCTGNACCCCATCAGAGCCTGGACACCTGGACCGAGTTTGTGGAGGAAA 1560
|||||
QY 1561 gaataagctgtgagcaagctgtgatccagaatgccaaactgtgtctgacctgtacaagt 1620
|||||
DB 1561 GAATAAGACTGTGAGCAAGCTGTGTATCCAGATGCCAATGCCACGTGTCTGCCATGTACAAGTG 1620
|||||
QY 1621 tglggtctcaacaagctgagccagatagagcctcatctactctatctatgtgaccacat 1680
|||||
DB 1621 TGTGGCTCTCCAAAGTGGCGCCAGGATGAGCGGCTCATCTACTTCTATGTGACCAACCAT 1680
|||||
QY 1681 cccgcagcggcttcaccatcgaatccaaagccatccgagggagctactagagggccagcgg 1740
|||||
DB 1681 CCGCGACGGCTTCACCATCGAATCCAAAGCCATCCGAGGAGCTACTAGAGGGCCAGCCGGT 1740
|||||
QY 1741 gctcctgagctgccaagccgaagctacaagtacgagctatcgtcgtgtacgcgtctcaa 1800
|||||
DB 1741 GCTCCTGAGCTGCCAAGCCAGCAGCTACAAAGTACGAGCATCTGCGTGTGTACCGGCTCAA 1800
|||||
QY 1801 cctgtccacgctgcacgatacgacgggaacccgctctgtcgcactgcaagaacgtgca 1860
|||||
DB 1801 CCTGTCCACGCTGCACGATGCGCACGGGAACCGCTTCTGTCTGACTGCCAAACCTGCCA 1860
|||||
QY 1861 tctgttcgcaacccctctgcccagctcagcggctgagggaggtggccacttggggcgccacgc 1920
|||||
DB 1861 TCTGTTCGCCACCCCTCTGGCCGCCAGCCTGGAGGAGTGGCACCTTGGGGCGCGCACGC 1920
|||||
QY 1921 cagcctcagcctgagtatcccccgctgcgcgcagcagcagggccactatgtgtgca 1980
|||||
DB 1921 CACGCTCAGCCTGAGTATCCCCGCGCTCGCCCGCCGAGCAGAGGCCACATGTGTGTGCGA 1980
|||||
QY 1981 agtgcgaagaccggcgagccatgacaagcactgccacaagaagtacctgtcgtgcaagc 2040
|||||
DB 1981 AGTGCGAAGACCGCGCAGGCATTCACAAGCACTGCCACAAGAAGTACCTGTCTGGTGCAGGC 2040
|||||
QY 2041 cctggaagccctcggctcacgacaacttgaccacactcctggtgaagcgtgagcactc 2100
|||||
DB 2041 CCTGGAAAGCCCTCGCTCAGCAGAACTTGACCAGCCTCTCTGGTGAACGTGAGCGACTC 2100
|||||
QY 2101 gctggagatgcaagtgttggcggagcgacgcgcccagcactcgtgtgtgtaacaaga 2160
|||||
DB 2101 GCTGGAGATGCAGTGTTGGTGGCGGAGCGCACGCGCGCCAGCATCGTGTGTACAAAGA 2160
|||||
QY 2161 cgaagggctgctggaagaaagtctggaagcgaacttggagctggaagcactcaaccagaagctgag 2220
|||||
DB 2161 CGAGAGCTGCTGGAGAAAGTCTGGAGTCCACTTGGCGGACTCCAAACCAAGAGCTGAG 2220
|||||
QY 2221 catccagcgtgctgcgaggaagatgcggagcgtatctgtgacgctgtgcaacgcaaa 2280
|||||
DB 2221 CATCCAGCGCTGCGGAGGAGGATGCGGAGCGTATCTGTGACGCGTGTGCAAGGCCAA 2280
|||||
QY 2281 gggctcgtcaactcctccagcagcgtggcgctggaagctccgaggaataaggcagcat 2340
|||||
DB 2281 GGGCTCGCTCAACTCTCCCGCAGCGTGGCGGTGGAAGGCTCCGAGGATAAGGCGAGCAT 2340
|||||
QY 2341 ggagatcgtgatccttctgtggtaccggcgctcagcgtgtcttcttctgtgggtccctcctc 2400
|||||
DB 2341 GGAGATCGTGATCTTGTGGTACCGGCTCATCGCTCTCTTCTCTGGGTCTCTCTCT 2400
|||||
QY 2401 cctcattctctgaacatgagagggccggccacgcagacatcaagcgggctacctgtc 2460
|||||
DB 2401 CCTCATCTTCTGTAACTGAGAGGCGCGCCACGACACATCAAGAGCGGCTACCTGTC 2460
|||||
QY 2461 catcatcatggccccgggaggtgcctcctggagagcaatgcgaataacctgtcctacga 2520
|||||
DB 2461 CATCATATGGACCCCGGGAGGTGCCTCTGGAGGAGCAATGCCAATACCTGTCTCTACGA 2520
|||||

QY 2521 tgcagccagtgaggaaattccccccgagagcggctgcacactggggaagtgctcggctacgg 2580
|||||
DB 2521 TCCACGCCAGTGGGAATTTCCCCCGAGAGCGGCTGCACCTGGGGAGAGTGTCTCGGCTACGG 2580
|||||
QY 2581 cgccttcgggaaggtggtggaagcctccgcttctcggaatccacaaggcgacgctgga 2640
|||||
DB 2581 GCCTTCCTGGGAAGTGGTGGAAAGCTCCGCTTTCGGCATCCACAAGGCGACGCTGTGA 2640
|||||
QY 2641 caccgtggcgtgaaaaatgctgaaagaggcgccacggccagcaagcaccgcgcgtgat 2700
|||||
DB 2641 CACCGTGGCGGTGAAAATGCTGAAAGAGGGCGCCACGGCCAGCAGCACCGGCGCTGAT 2700
|||||
QY 2701 gtcgagctcaagatcctcatctcaatcggcaacacactcaacgtggltcaacctcctcgg 2760
|||||
DB 2701 GTCGGAGCTCAAGATTCCTTCATTCATCGGCAACCACTCAACGTGGTCAACCTCCTCTGG 2760
|||||
QY 2761 ggcgtgcaccagccgagggcccccctcatgtgatcgtggagttctgcaagtacggcaa 2820
|||||
DB 2761 GCGTGTCAACCAAGCGCAGGGCCCTCATGTGTATCTGTGAGTCTGTCAAGTACGGCAA 2820
|||||
QY 2821 cctctccaaacttctcgtgcgccaagcgagcgtcttcaagccctgcgcggagaaagtctcc 2880
|||||
DB 2821 CCTCTCCAATTCCTTGC CGCAAGCGGGACGCTTTCAGCCCCCTCGCGGAGAACTCTCC 2880
|||||
QY 2881 cgaacagcgagcagccttccgcgcgcagctggtgagctcgcaggtgagctcggagcggcc 2940
|||||
DB 2881 CACGAGCGCGGAGCGCTTTCGCGGCCATGTGTGAGCTCGCCAGGCTGGATCGAGAGCGGCC 2940
|||||
QY 2941 gggagcagcgacaggggtctcctctcgcgcggttctcgaagaccgagggcgagcgagcg 3000
|||||
DB 2941 GGGGAGCAGCAGGAGGTCTCTTTCGCGCGGTTCTCGAAGACCAGGCGGAGGCGAGGCG 3000
|||||
QY 3001 ggccttctccagacaagaagctgaggacctgtggtcgtgagcccgctgacctggaagatct 3060
|||||
DB 3001 GGCCTTCTCCAGACCAAGCTGAGGACCTGTGTGCTGAGCCGCTGACCATGGAAGATCT 3060
|||||
QY 3061 tctcgtctacagcttccaggtgagcagaggtgaggttctcgtgcttcccgaagatcat 3120
|||||
DB 3061 TGTCTGTACAGCTTCCAGGTGGCCACGAGGATGAGGTTCTGTGCTTCCGCAAGTGCAT 3120
|||||
QY 3121 ccacagagacctggtcgtcgtcggaaacatctcgtcgtcggaaagcgactggtgaagatctg 3180
|||||
DB 3121 CCACAGAGCTGTGCTCTCGGAACATTCCTGTGCGAAGACGACCTGTGTGAAGATCTG 3180
|||||
QY 3181 tgccttggccttcccgggacatctcaagaccctgactacgtccgcaaggcagctgc 3240
|||||
DB 3181 TGACTTGGCTTCCCGCGGACATCTACAAGACCCCTGACTACTCTCCCAAGGCGACTGC 3240
|||||
QY 3241 ccggtgccctgaaatggaatggccctgaaagcatcttcgacaaggtgtacaccacgca 3300
|||||
DB 3241 CCGGCTGCCCTGAAGTGGATGGCTGCGCCCTGAAGCATCTTCGACAAGGTGTACACCGCA 3300
|||||
QY 3301 gagtgcgtgtgctccttgggtgcttctcgtggagatcttctcgtggggcctcccc 3360
|||||
DB 3301 GAGTGACGTGTGTCTTTTGGGGTGTCTCTGTGGAGATCTTCTCTGTGGGGCTCTCCC 3360
|||||
QY 3361 gtacctggggtgcagatcaatgagggattctgcagcggctgagagacggcgcaagat 3420
|||||
DB 3361 GTACCTTGGGTGCAGATCAATGAGGAGTTCGCCAGCGCTGAGAGCGGCACAGGAT 3420
|||||
QY 3421 gaggccccgagctggcactcccccatcagccgcatctgctgaactgctggtccgg 3480
|||||
DB 3421 GAGGGCCCCGAGCTGGCCACTCCCCCATACGCCCATCATGCTGAACCTGTCTGTCGG 3480
|||||
QY 3481 agaccaccaagcgagacctcattctcggagctggtggagatccttggggacctgtcca 3540
|||||
DB 3481 AGACCCCAAGCGCAGACCTGCAATCTCGGAGCTGTGTGGAGATCTCTGGGGACCTGTCTCA 3540
|||||
QY 3541 gggcagggggcctcgaagaggaagaggtctgcatggtccccgcgcagctctcagagctc 3600
|||||
DB 3541 GCGAGGGGGCTTCAAGAGAGAGAGGAGGTCTCATGGGCCCGCGCAGCTCTCAGAGCTC 3600
|||||
QY 3601 agaagagggcagcttctcgcagggtgltccaccatggccctacacatcgccccaggtcagc 3660
|||||

QY 3061 tgctgtacagcttcaggtgcccagagagatggagttctctggttccgaaagtgcac 3120
Db 3061 TGCTGTCTACAGCTTCCAGTGGCCAGAGAGATGGAGTTCTTGGCTTCCGAAAGTGCAT 3120
QY 3121 ccacagagactggctgtcggaaacattctgtctcggaaagcagctggtgaagatctg 3180
Db 3121 CCACAGAGACTGGCTGCTCGGAACATTTGCTGTGGAAAGGAGCTGTTGAAGATCTG 3180
QY 3181 tgactttggctgtcccgaggacatctacaaagacctgactacgtccgcgaagggcagtg 3240
Db 3181 TGACTTTGGCTTGGCCGGGACATCTACAAAGACCTGTACTCGCAAGGCGCAGTGC 3240
QY 3241 ccgggtccctgaagtggatggccctgaagacatcttcgacaaggtgtacacacagca 3300
Db 3241 CCGGCTCCCTGAAGTGGATGGCCCTGAAAGCATCTTCGACAAGGTGTACACACAGCA 3300
QY 3301 gagtacgtgtggtcccttgggtgctctctgtggagatctctctctggggcctccc 3360
Db 3301 GAGTACGTGTGTTGGGTGCTTCTCTGGGAGATCTTCTCTGGGGGCTCCCT 3360
QY 3361 gtacctgggtgcagatcaatgagggtctctccagcggctgagagacggcacaagat 3420
Db 3361 GTACCTTGGGTGCAGATCAATGAGGATTTCTGCCAGGCTGTGAGAGCGCACAAAGAT 3420
QY 3421 gaggcccgagctgacctcccgccatccgacccgcatcatgctgaactggtgctcgg 3480
Db 3421 GAGGCCCCGGAGCTGGCCACTCCCGCCATACGCCGATCATGCTGAACCTCTGCTCGG 3480
QY 3481 agaccccaagcggagactgctctctggagctggtggagatcctggggacctgctcca 3540
Db 3481 AGACCCCAAGCGAGACTGCTATCTCGGAGCTGTGGAGATCTTGGGGACCTGTCTCA 3540
QY 3541 gggcagggcctcaagaaagaggggtctgcatgcccccgccgctcagagctc 3600
Db 3541 GGGCAGGGGCTGCAAGAGGAGGAGGCTGCTGATGCCCCCGCCAGCTCTCAGAGCTC 3600
QY 3601 agaaagggcagcttctgcaggtgtccacatggccctacacatcgccaggctgacgc 3660
Db 3601 AGAAGAGGCGAGCTTCTCGCAGGTGTCCACCATGGCTTACACATCGCCCGCAGCTGACGC 3660
QY 3661 tgagagaccccgcaagctgagccacagcctgagccaggtattacaactgggt 3720
Db 3661 TGAGGACAGCCCGCAAGCCTGCGAGCCAGCCTGCGCCAGAGTATTACAACTGGGT 3720
QY 3721 gtccttcccggtgctgcccagagggctgagaccgtggttctccagatgaagac 3780
Db 3721 GTCCTTTCCCGGTGCTTGGCCAGAGGGGCTGAGACCGTGTCTTCTCCAGGATGAAGAC 3780
QY 3781 attgaggaattcccatgaccccaagacctacaaaggctctgtgacaacagagaga 3840
Db 3781 ATTGAGGAATTTCCCATGACCCCAACGACCTACAAAGGCTCTGTGGACAACCCAGACAGA 3840
QY 3841 cagtgagatggtgctggcctcgaggagtttgagcagatagagcagggcagatagacaaga 3900
Db 3841 CAGTGGATGTTGCTGGCTCGGAGAGTGTGAGCAGATAGAGCAGGATAGACAAGA 3900
QY 3901 aagcggcttcagct 3914
Db 3901 AAGCGGCTTTCAGGT 3914

RESULT 4
US-08-901-710-1
; Sequence 1, Application US/08901710
; Patent No. 6107046
; GENERAL INFORMATION:
; APPLICANT: Alitalo, Kari
; APPLICANT: Aprelikova, Olga
; APPLICANT: Pajusola, Katri
; APPLICANT: Armstrong, Elina
; APPLICANT: Korhonen, Jaana
; APPLICANT: Kaipainen, Arja
; APPLICANT: Mälikäinen, Marja-Terttu

; TITLE OF INVENTION: FLT4, A RECEPTOR TYROSINE KINASE, AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/901,710
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,011
; FILING DATE: 14-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,754
; FILING DATE: 09-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/959,951
; FILING DATE: 09-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28113/33824
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 20..3916
; US-08-901-710-1

Query March 95.2% Score 3912.4; DB 3; Length 4195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3913; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ccacgcagcggcgagatgcagcggggcgccgctgctgctgagctggtgctg 60
Db 1 CCACGCGCAGCGCGCGAGATGCAGCGGGCGCGCGCTGTGCTGCGACTGTGCTG 60
QY 61 cctgagactcctgagcggcctggtgagtgactcctacacccccccgacctgaaacat 120
Db 61 CCTGGAGACTCTTGACCGCCCTGGTGGTGGCTACTCCATGACCCCCCGACCTTGAACAT 120
QY 121 cagcggaggtcacacatcatcgacaccggtgacagcctgtccatctcctgcaggggaca 180
Db 121 CAGCGAGAGTCAACACGTCATCGACACCGGTGACAGCTGTCCATCTCTCTGACGGGACA 180
QY 181 gacccccctcagtgggctggccagagagctcagagggcgccacccacagacaagga 240
Db 181 GCACCCCTCGAGTGGGCTTGGCCAGGAGCTCAGAGGGCGCCAGCCACCCGAGACAAGGA 240
QY 241 cagcggagacacgggggtggtgagagactgcaggggcacagacgcccagggcctactgcaa 300
Db 241 CAGCGAGGACACGGGGTGGTGGGAGAGCTGCGAGAGTGGGAGGCGCACAGACGCCAGGCCCTACTGCAA 300

QY 301 ggtgttgctgcagcagggtacatgccaacgacacacaggaagcagctacgtctgtactacaa 360
DB 301 GGTGTGTGCTGTCACGAGGTATATGCGCAACGACACAGGCAGCTACGCTCTACTACAA 360
QY 361 gLacatcaaggcaagcatcgaggccaccacgcccagcagctccctacgtgttcgtgagaga 420
DB 361 GTACATCAAGGACACCATCGAGGCGACACCGCCCGCAGCTTCTACGTTGCTTGACAGA 420
QY 421 ctttgagcagccattcatcaaaagctgacacgctcttggltcaacaggaaggcgcacat 480
DB 421 CTTTGAGCAGCCATTATCAAAAGCCGTGACACGCTCTTGGTCAACAGGAAGGACGCCAT 480
QY 481 ggggtgcccctgtcgtgtccatcccccgccctcaatgtcaagctgagctgcgcaaaagctc 540
DB 481 GTGGGTGCCCCTGTCTGGGTGTCATCCCGGCTCAATGTACGCTGGCGCTCGCAAAAGCTC 540
QY 541 ggtcgtgtggccagagggcaggaggtggtgtgggataaccgagcgggagcatgctcgtgtc 600
DB 541 GGTGCTGTGGCCAGAGCGGAGGAGGTGCTGTGGGATGACCGCGCGGSCATGCTCGTGTC 600
QY 601 cagcgcactgctgcagatgcccctgtacctgagtgagagaccacctggggagaccagga 660
DB 601 CACGCGACTGCTGCAGATGCGCTGTACCTGCAGTGGAGACCACCTGGGAGACCAGGA 660
QY 661 cttctttccaaaccttccctggtgacatcacaggaacagagctctatgacatccagct 720
DB 661 CTTCCTTTCCAAACCCCTTCTGCTGTGTGCATCACAGGAACAGGCTCTATGATCATCCAGCT 720
QY 721 gttgccaggaagtcgctggagctgctggtgagggagagctggttcctgaactgcacgct 780
DB 721 GTTGCCAGGAAGTCCGTGGAGCTGCTGCTAGGGGAGAACTGCTTGAACATGCACCGT 780
QY 781 ggggctgaatttaactcaggtgtcaacctttagctgggactaccacggaggaagcagga 840
DB 781 GTGGGCTGAGTTTAATCACTCAGGTGTACCTTTGACTGGGACTACCCAGGAAAGCAGGAGA 840
QY 841 gcggggttaagtgggtgcccagcagcagctcccagcagacccacacagaactctccagcat 900
DB 841 GCGGGGTAAGTGGGTGCCCCAGGACGCTCCACAGACAGCCACACAGACTCTCCAGCAT 900
QY 901 cctgaccatcccaacgctcagccagcagcagctgggctcgtatgtlgtcaaggccaacaa 960
DB 901 CCTGACCATTCCAAACAGTCAGCCAGCAGCAGCTGGGCTCGTATGTGTGCAAGGCCAACAA 960
QY 961 cggcatcagcagatttcgggagagcagcaggtcatgtgcatgataaaatcccttcatcag 1020
DB 961 CGGCATCCAGCATTTTCGGGAGAGCACCAGGTCATTGTGTCATGAAAATCCCTTCATCAG 1020
QY 1021 cgtcaggtgctcaaaggaccatctctggagccagcagcagcagcagcagcagcagcagc 1080
DB 1021 CGTCAGTGGCTCAAAAGGACCATCTCGAGGCCACGCGCAGAGACGAGCTGTTGAAGCT 1080
QY 1081 gccctgaaagctggcagcgtaccctcccgagctccagtggttacaaggagcaagggc 1140
DB 1081 GCCCGGTAAAGCTGGCAGGCTACCCCGCCGCGAGTTCCAGTGGTACAAGGATGGAAGGC 1140
QY 1141 actgtccgggcccacagtcacatgcccctggtcgtcaagagatgacagagggccaacac 1200
DB 1141 ACTGTCCGGGCCCACAGTCCACATGCCCTGCTCTCAAGGAGGTGACAGAGGCCAGGCAC 1200
QY 1201 aggcacctacacctcgcctgtggaactccgctgctggcctgagggcgcaacatcagcct 1260
DB 1201 AGGCACCTACACCCCTGCGCCCTGTGGAAGTCCGCTGCTGGCCCTGAGGCGCAACATCAGCCT 1260
QY 1261 ggaagctggtggaatgtccccccacagatatacagaagagagcccttccctcccagcat 1320
DB 1261 GGAGCTGGTGGTGAATGTGCCCCCCCCAGATACATGAGAAGGAGGCGCTTCTCCCCCAGCAT 1320
QY 1321 ctactcgttcacagcccgccagccctcaactgcacggcctacggggtgcccctgcctct 1380
DB 1321 CTACTCGGCTCACAGCCGCCAGGCCCTCACCTGCAGGSCCTACGGGTGCCCCCTGCCCTCT 1380
QY 1381 cagcatccagtgagcagggccctggaacacctggaagatggtttggcccagcgtagctct 1440

DB 1381 CAGATCCAGTGGGCATGGCGGCCCTGGACACCTTGAAGATGTTGCCAGCGTACTCT 1440
QY 1441 ccgscggcggcagcagcaagaacctcatgccacagtgccgtgactggagggcggtgaccac 1500
DB 1441 CCGCGCGCGGCAGCAGCAAGACCTCATGCCACAGTGGCGGTGACTGGAGGGCGGTGACCAC 1500
QY 1501 gcaggaTgcctggaaccccccatcgagagcctggacacctggacagagtttggaggagaa 1560
DB 1501 GCAGGATGCGGTGAACCCCATCGAGAGCTGGACACCTTGGACCGAGTTTGTGGAGGAAA 1560
QY 1561 gaataaagactgagcaagcgtggtgatccagaatgcaacgctgtctccatgtataaagt 1620
DB 1561 CAATAAGACTGTGAGCAAGCTGGTGATCCAGATGCCAAGCTGTCTGCCATGTATAAGTG 1620
QY 1621 tglqatctccaacaagylggcagaggtgagcgctcatctacttctatgtgaccacat 1680
DB 1621 TGTGCTCTCCAACAAGTGGGCCAGGATGAGCGGCTCATCTACTTCTATGTGACCCACAT 1680
QY 1681 cccgcagcgttccaccatcgaatccaagccaatccgagagctactagaggccagccggt 1740
DB 1681 CCGCGAGGGCTTCAACATCGAATCCAAGCCATCGAGGAGCTACTAGAGGGCCAGCCGGT 1740
QY 1741 gctcctgagctgcaagcgaacagctacaagtacagacatctcgctgagctgcaagaagctgca 1800
DB 1741 GCTCCTGAGCTGCCAAGCCGACAGCTACAAGTACGACATCTGCGCTGGTACCOCCTCAA 1800
QY 1801 cctgtccacgctgcacgatgcgcaggggaacccgcttctgctgactgcaagaagctgca 1860
DB 1801 CCTGTCCACGCTGCACGAATCGGCACGGGAACCCGCTTCTGCTCGACTGCAAGAAGCTGCA 1860
QY 1861 tctgttcycaacctctggccgcagcctcgagagagtgagcactggcgagccagcgc 1920
DB 1861 TCTGTTGCCACCCCTCTGGCCGCGCAGCTGGAGGAGTGGCAGCTGGGCGGCCACGC 1920
QY 1921 cacgctcagctgagtatcccccgctgcgcgcagcagcagcagcagcagcagcagcagc 1980
DB 1921 CACGCATCAGCTGAGTATCCCCGCTGCGCCGCGCAGCAGCAGGCGCCACTATGTGTGCGA 1980
QY 1981 agtcaagaagcggcgcagccatgacaagcagctgccaaagctacctgtcggtgagggc 2040
DB 1981 AGTCAAGAAGCGCGCGCAGCCATGACAAGCACTGCCACAAGAAGTACCTGTCGGTGCAGGC 2040
QY 2041 cctggaagccctcgtcgtcagcaagaactgacagacactcctcgtgagcgtgagcagc 2100
DB 2041 CCTGGAAGCCCTCTCGGCTCAGCGAGAACTTGACCGACTCTCTGGTGAACGTGAGCGACTC 2100
QY 2101 gctgagatgagtgctgtggtggcggagcgacgcgcccagcagctcgtgtggtacaaga 2160
DB 2101 GCTGGAATGCAGTGTCTGTTGGTGGCGGAGCGCACGCGCCGCGCAGCATCTGTGGTACAAGA 2160
QY 2161 cgaagtgctgtagagaaaagtctggagtcgactggcgagctcccaaccagaagctgag 2220
DB 2161 CGAGAGGCTGCTGGAGGAAAAGTCTGGAGTGCAGTGGCGGACTTCCAACCAAGAGCTGAG 2220
QY 2221 catccagcgtgcgcagagagatgcggagcgtactctgtgagcgtgtgcaacgcaa 2280
DB 2221 CATCCAGCGGTGCGCGAGAGGATGCGGAGCGTACTGTGTGACGCTGTGCAACGCCAA 2280
QY 2281 gggctgcgtcaactccctcccgccagcgtggcggtagaggctcccgagataaggagcagcat 2340
DB 2281 GGCGTGCATGCTTCCGCTCCGCGAGCTGGCGTGGCAAGGCTCCGAGCATTAAGGCGACAT 2340
QY 2341 ggaagatcgtactccttctcgggtaccggtcactcgtgtcttcttctgggtcctcctcct 2400
DB 2341 GGAGATCGTATCTCTGCTACCGCGGTGCTGCTGTCTTCTTCTGGGTCTCTCTCT 2400
QY 2401 cctcactctctgataacatgaggagggcgcccaacgagacatcaagcggggtacctgtc 2460
DB 2401 CCTATCTTCTGTAAATGAGGAGGCGGCCACGCGACATCAAGAGGCTACCTGTCT 2460
QY 2461 catcatcatggagccccgggaggtgcctctggagagcaatgcgaatacctgtcctacga 2520

|||||
Db 1452 CCGCGCGGCGAGCAGCAAGACCTCATGCCACAGTGGCGTACCTGGAGGGCGGTGACCCAC 1511
Qy 1501 gcaagatgcccgtgaaccccaatcgagagcctggacaacactggaccgagttgttgaggagaaa 1560
|||||
Db 1512 GCAGGATGCCGTGAACCCCATCGAGAGCTGGACACCTGGACCGAGTTGTGGAGGAAA 1571
Qy 1561 qaatgaagactgagcaagctggatgagccagaatgcaaacgtgctcgcattgacaagt 1620
|||||
Db 1572 GAATAGACTGTGACAAAGCTGGTATCCAAATGCCAACTGTCTGCCATGTACAAGTG 1631
Qy 1621 tgggtctcaacaagggtggccaggatgagcgctcatctacttctatgtgacaccat 1680
|||||
Db 1632 TGTGTTCTCCAAGAGTGGCGAGGATGAGCGGCTCATCTACTTCTATGTGAGCACCAT 1691
Qy 1681 ccccgagcgttcacattcgaatcaaaccaatccagagagctactaagagccagcaggt 1740
Db 1692 CCGCGAGCGGTTTACCATCGAATCCAAAGCCATCCGAGGAGCTACTAGAGGGCCAGCCGGT 1751
Qy 1741 gctcctgagctccaagccgacagctacaagtacgagcatctggcgtggtaccgcctcaa 1800
Db 1752 GCTCTGAGTGCACAGCGCAGACTACAAGTACGAGCATCTGGCTGTGTACCGCTCAA 1811
Qy 1801 cctgtccagctgacagatgacagcagcaggaacccgctctctgctcagctgcaagaacgtca 1860
|||||
Db 1812 CCTGTCCAGCTGCAGATGCGCAGCGGAACCGCTTCTGCTCGACTGCAAGAACGTGCA 1871
Qy 1861 tctgttcgcaacccctctgtgcccagcctgagagagtgagcactggcgccgacgc 1920
Db 1872 TCTGTTTGGCCACCCCTCTGGCCGCCAGCTTGAGAGGTGGACCTTGGGGCCGCCACGC 1931
Qy 1921 cagcgtcagcctgagtatcccccgctgcgcccagcagcagggccactatgtgcga 1980
Db 1932 CAGGCTCAGCCTGAGTATCCCGCGCTGCGCCGAGCAGCAGGCGCCACTATGTGTGCGA 1991
Qy 1981 agtgaagaccgagcagcactgacaagcactgcaagaagtagctcgtgtagcagc 2040
Db 1992 AGTGAAGACCGCGCAGCCATGACAAGCAGCTGCCACAAAGTAGTACCTGTCTGGTCAAGC 2051
Qy 2041 cctggaagccctcgtgctcagcagaaacttgaccgacctcctggtgaactgagcactc 2100
Db 2052 CCTGGAAGCCCTCGGCTCAGCAGAACTTGACCGACCTCTCTGTTGAACGTGAGCGACTC 2111
Qy 2101 gctgagagatgcaagtgcttggtgcccagcagcagcagcccaacatcgtggtacaaga 2160
Db 2112 GCTGAGATGCAAGTGTGTTGGTGGCGAGCGCAGCGCCGCGGACATCGTGTGTACAAGA 2171
Qy 2161 cgaagagctgctgaggaagaaagtctggaagtcgaactggcgagactccaacagaaagtga 2220
Db 2172 CGAGAGCTGCTGGAGGAAAAGTCTGGAGTGGACTTGGCGGACTCCAACACAGAGTGAG 2231
Qy 2221 catccagcgtgctgcgagagagatgcygagcgtatctgtgagcgtgtgcaacgcaa 2280
Db 2232 CATCCAGCGCTGCGGAGGAGATGCGGAGCGCTATCTGTGCAGCGTGTGCAACGCCAA 2291
Qy 2281 ggcgtcgtaactcctccgacagctgagccatgaaagctccagagataaaggaagat 2340
Db 2292 GGGCTCGCTCAACTCTCCGACAGCTGGCGGTGAAGGCTCCGAGGATAAGGGACCAT 2351
Qy 2341 ggaagctgtagctctgctggtaccggtcagctcgtctctctctggtgctcctcct 2400
Db 2352 GGAGATCGTATCTTGTTCGGTACCGGCTCATCGCTGTCTTCTTCTGGGTCTCTCTCT 2411
Qy 2401 cctcatctctgtaacatgagagagccgcccacgcagacatcaagagcgggttaacctgctc 2460
Db 2412 CCTCATCTTCTGAACATGAGGAGCGCGCCACCGCAGACATCAAGACGGGTACTCTGTC 2471
Qy 2461 catcatatgaccccgagaggtgctctgagagagcaatgcgaataacctgctcactca 2520
Db 2472 CATCATGAGACCCCGGAGGAGTGGCTCTGGAGGAGCAATCGGAATACCTGTCTCTACCA 2531
Qy 2521 tgcagcagtggaattcccccgagagcggctgacactgagggagagtgctcgtcagcag 2580
|||||

Db 2532 TGCCAGCAGCTGGAAATTCCTCCCGAGAGCGCTGCACCTGGGGAGAGTGTCTGGCTACGG 2591
Qy 2581 cgccttcgggaaggtggtggaagcctccgcttcgagcatccacaagggcagcagctgtga 2640
|||||
Db 2592 CCGCTTCGGGAAGGTGTGGAAGCTCCGCTTTCGGCATCCACAAGGCGACGACTGTGA 2651
Qy 2641 caccctggcgtgaaatgctgaaagagggcgccacggccagcagcagcagcagcagcagc 2700
|||||
Db 2652 CACCCTGGCCCTGAAATGCTGAAAGAGGCGCCACGGCCAGCGACACCGCGCGCTGAT 2711
Qy 2701 gtcgaggtcaagatcctcattcaatcgcgaacacactcaactggttcaacctcctcgg 2760
Db 2712 GTGCGAGCTCAAGATCTCTATTACATCGCAACACCTCAACGTGTGTAACCTCTCTCGG 2771
Qy 2761 ggcgtgcaccaagccgagggccccctcatggtgagtggtgagttctgcaagtagcggcaa 2820
Db 2772 GCGGCGCACCAGCGCAGGGCCCCCTCATGCTGATCTGTGAGTTCTGCAAGTAGCGCAA 2831
Qy 2821 cctctccaactcctgctgagcgaagcggcagccttcagccccctgcgaggaagatctcc 2880
Db 2832 CCTCTCCAACCTTCTTGGCGGCAAGCGGAGCGCTTCAAGCCCTTCGCGGAGAACTCTCC 2891
Qy 2881 cgaacagcgcgagcagcctcccgccatggtgagcctcgcaggtcggatcgagagcgccc 2940
Db 2892 CGAGCAGCGCGGAGCGCTTCCGCGCCATGTTGGAGCTCGCCAGGCTGGATCGGAGCGGCGC 2951
Qy 2941 gggagcagcagcagcagcctctctcgcgcttctcgaagaccgagggcgagcagcagc 3000
Db 2952 GGGGAGCAGCGACAGGGTCTTTCGCGCGGTTCTCGAAGACCGAGGCGGAGCGAGCG 3011
Qy 3001 ggcctctccagacaagaagctgaggaacctgtggctgagccccctgacacttggaagatct 3060
Db 3012 GGCTTCTCCAGACCAAGAAAGCTGTGGCTGAGCCCCGTGACCATGGAAGATCT 3071
Qy 3061 tgcctgtacagcttccaggtggccagaggtgaggttccctggcttccccgaaagtgc 3120
Db 3072 TGTCTGTACAGCTTCCAGGTGGCCAGAGGATGAGTCTCTTGGCTTCCCGAAGTGCAT 3131
Qy 3121 ccacagagacctgctgctcggaaacttctgctgcgaaagcagcgtggtggaagatctg 3180
Db 3132 CCACAGACCTTGGCTGCTCGGAACATTCGCTGTCGAAAGCGACGTGGTGAAGATCTG 3191
Qy 3181 lgaactlgaactlcccgagacatctacaagacctgaactacgtccgaaagggcagtc 3240
Db 3192 TGACITTGCTTGGCTTGGCCGGACATCTACAAAGACCTGACTACTCTCCGCAAGGCGAGTGC 3251
Qy 3241 ccggctgccccctgaagtggatggccccctgaaagcatcttcgacaaggtgtacacacgca 3300
Db 3252 CCGGCTGGCCCTGAAGTGGATGGCCCTTGAAGACATCTTCGACAAGGTGTACACCGCA 3311
Qy 3301 gagtgaactggtgctccttgggtgctctctctggagatcttctctgggggctcccc 3360
Db 3312 GAGTACGCTGGTCCCTTGGGGTGTCTCTGAGAGATCTTCTCTGAGGGGCTCTCC 3371
Qy 3361 gtacctgggtgagacatcaatgagaggttctgcagcggtgagagcggcagcagat 3420
Db 3372 GTACCTCTGGGTGCAGATCAATGAGAGTTCCTGCCAGCGCTGAGAGCGGCAAGGAT 3431
Qy 3421 gaggccccggagctggcaactccccccatacgcgcgcatcatgtgtaactgctggtccgg 3480
Db 3432 GAGGCGCGCGAGCTGGCCACTCCCGCCATAGCGCGCATCATGCTGAACTGTGTGTCGG 3491
Qy 3481 aqaccccaagcagcagcctgcatctcgagctggtgagatctcgggggacctctcca 3540
Db 3492 AGACCCCAAGCCAGCACTGTCATTTCTCGAGTTCCTCGAGCTGGTGGAGATCTTGGGGACCTCTCTCA 3551
Qy 3541 gggcaggggcttgcgaaggaagagaggtctgtcagctggccccgcagcagctctcagagctc 3600
Db 3552 GGGCAAGGGCTTGAAGAGGAAGAGGAGGTCTGCATGGCGCCCGCGCAGCTCTCAGAGCTC 3611
Qy 3601 aqaaagggcagccttctcgcaggtgtccacctggcccttaacatcgcccaagctgagc 3660
Db 3612 AGAAGAGGCGAGCTTCTCGCAGGTGTCCACCATGGCCCTTACACATCGGCCCGGCTGACGC 3671

QY 1081 gccctggaagctggcagctgaccccccgcgcgaattccagtggttacaaggatgaaaggc 1140
DB 1092 GCCCGTGAAGCTGGCAGCGGTACCCCGCCCGCGAGTTCCAGTGGTACAAGGATGAAGGC 1151
QY 1141 actgtccgggggcacacagtcacacatgccttgctcgaagaggtgacagggccagcac 1200
DB 1152 ACTGTCCGGCGCCACAGTCCACATGCCCTGGTCTCTCAAGGAGGTGACAGAGGCCAGCAC 1211
QY 1201 aggcacctacacctgccttgcttggaactccgctgctgctgctgctgctgctgctgctgct 1260
DB 1212 AGGCACCTACACCTCGCCCTGTGGAACTCCGCTGCTGGCTTGAGGCCCAACATCAGCCT 1271
QY 1261 ggaagctgggtggaatgtyccccccacagatataatgagaagagggcctcctccccagcat 1320
DB 1272 GGAGCTGGTGGTGAATGTGCCCGCCCGCCAGATACATGAGAAGGAGGCCCTCTCCCGCCAGCAT 1331
QY 1321 ctacteggtcacagccgcccagggcctcacctgcagcgcctacgggtgcccctgctcct 1380
DB 1332 CTACTCGGCTCACAGCGCGCAGGCCCTCACCTGCACGGCCTACGGGCTGCCCTGCCCT 1391
QY 1381 caagctccagtgacactggcgcctggacacccctgacagatgtttgcccagcgtagctct 1440
DB 1392 CAGCATCCAGTGGCACTGGCGGCCCTGGACACCCCTGCAAGATGTTGCCACGCTAGTCT 1451
QY 1441 ccggcgggcgagcgaagaaacctcatgcccagtgcccgctgactggagggcggtgacacac 1500
DB 1452 CCGCGCGCGCAGCAGCAAGACCTCATGCGCACAGTGGCTGACTGGAGGGCGGTGACCCAC 1511
QY 1501 gcaagatgcgtgaaacccatcgagagcctggacacccctggaagatgtttgagagggaaa 1560
DB 1512 GCAGGATGCGGTGAACCCCATCGAGAGCTGGACACCTGGACCGAGTGTGTGAGCGAAA 1571
QY 1561 gaataagactgtagcaagctggtgatccagaatgccaaagtgatgctgctgctgctgctgct 1620
DB 1572 GAATAAGACTGTGAGCAAGCTGTGATCCAGAAATGCCAACGTTGTGCCATGTACAAGTG 1631
QY 1621 tgtgtctccaaagtggtggcagatgagcgctcatctactctctatgtgacacacat 1680
DB 1632 TGTGCTCTCCACNAGTGGCGCAGGATGAGGGGCTCATCTACTTATGTGACACCCAT 1691
QY 1681 ccccgagcgttcacatccaatccaaagccatccgagagctactagagggccagccggt 1740
DB 1692 CCCCGAGCGGCTTCAACATCGAATCCAAAGCCATCCGAGGAGCTACTAGAGGGCAGCCGCT 1751
QY 1741 gctcctgagctccaaagccagacgttacaagttacaagcaltcgcgtggtaccgctcaa 1800
DB 1752 GCTCCTGAGCTGCCAAGCCAGCAGCTACAAGTACGAGCATCTGCCGTGCTACCGCTCAA 1811
QY 1801 cctgtccagctgcagatgacgacgggaacccgcttctgctgactgacgacgacgacgac 1860
DB 1812 CCGTCCACGCTGCAGATGCGCAGCGGAACCCGCTTCTGCTGCTGCTGCAAGAACGTGCA 1871
QY 1861 tctgttcgcaacccctctgagccagcctggaagaggtggaaccttgaggcgccacagc 1920
DB 1872 TCTGTTCCGACCCCTCTGCGCCCGCAGCCTGGAGAGGTGGCACCTTGGGGCGCCGACCGC 1931
QY 1921 cagctcagctgagtatcccccgctgcgcgccgagacgagggccactatgtgtgcga 1980
DB 1932 CACGCTCAGCCTGAGTATCCCGCGCTGCGCGCCGACGACGAGGGCCACTATGTGTGCGA 1991
QY 1981 agtgcaagccgagccatgacaagcactgccaagaagatcactgctgctgagc 2040
DB 1992 AGTGCAAGACCGGCGAGCCATGCAAGACTGCAAGACTGCCAAGAAGTACTCTGTGCTGCGAGGC 2051
QY 2041 cctggaagccctcgtgtcacgcagaaactgaccgacctcctggtgaaagctgagcactc 2100
DB 2052 CCTGGAAGCCCTCGGCTCACGAGAACTTGACCGACCTCCTGGTGAACGTGACGCGACT 2111
QY 2101 gctggaagatcagtgcttggtgcccggagcgcaagcgcacagcctcagctggtgtacaaa 2160
DB 2112 GCTGGAGATGAGTGTGTGTGGCGGAGCGCACGCGCCCGCAGCATCGTGTGGTACAAAGA 2171

QY 2161 cgaagggctgctggagaaagtctgagtcgacttggcggaactccaaaccagaagctgag 2220
DB 2172 CGAGAGGCTGCTGGAGGAAAAGTCTGGAGTCGACTTGGCGACTTCCAAACGAAGCTGAG 2231
QY 2221 catccaagcgtgcgcgagagagatgcgaacgctaltctgtcagcgtgtgcaacgcaa 2280
DB 2232 CATCCAGCGGTGGCGGAGGAGGATGGGGACGCTA TCTGTGACCGGTGTGCAACGCCAA 2291
QY 2281 gggctgctcaactcccgccagcgtggcgtggaaggtccgagagataagggcagcat 2340
DB 2292 GGGCTGCTCACTCCGCCAGCGTGGCGTGGAAAGGCTCCGAGGATAAAGGCGACAT 2351
QY 2341 ggaatcgtgatacctctcgtgaccggtcgcgtcgcgtcgcgtcgcgtcgcgtcgcgtcgc 2400
DB 2352 GGAGATGCTGATCCTTGTGGGTACCGCGCTCATCGCTGTCTTCTTCTGGGTCTCTCTCT 2411
QY 2401 cctcatctctgtaacatgagagagccgcccacgcagacatcaagacgggctacctgtc 2460
DB 2412 CCTCATCTTCTGTAACTAGAGAGCGCCGCCACGAGACATCAAGACGGGCTACCTGTCT 2471
QY 2461 catcatcatggaaccccgggagggtgcctctctggagagcaatgcgaataacctgtcctca 2520
DB 2472 CATCATCATGACCCCGGGAGGTGCCCTCTGGAGGAGCAATGCCAATAACCTGTCTTACGA 2531
QY 2521 tgcagcagctgggaattcccccgagagcgtgcacctggggagagtgtctcggctacgg 2580
DB 2532 TGCCAGCCAGTGGGAATTCCTCCCGAGAGCGGCTGCACCTGGGGAGAGTGTCTGGCTACGG 2591
QY 2581 cgcctccgggaaggtggtggaagcctccgtctcgggcataccacaaggagcagcgtgtga 2640
DB 2592 GCGCTTCGGGAAGGTGTGTGGAAGCTTCCGCTTTCGGATCCACAAAGGCGACAGCTGTGA 2651
QY 2641 caccgtgcccgtgaaatgctgaaagagcgccacggcagcagcagcagcagcagcagcagc 2700
DB 2652 CACCTGCGCCGTGAAAATGCTGAAAGAGGGCGCCACGGCCAGCGACACCGCGCTGAT 2711
QY 2701 gtcggagctcaagatcctcatctcaatcagtcggcaacacacctcaacgtggttcaacctctc 2760
DB 2712 GTCGGAGCTCAAGATCTCTATTCATCTCGCAACCACTCAACCTGTGTCAACCTCTCTCGG 2771
QY 2761 ggcatacacaagccgcagggcccccctcatgctgagtgagtgcttctgcaagtcagcga 2820
DB 2772 GCGCTGCAACCAAGCCGACGGCCCTCTCTGATCTGATCTGATCTGATCTGATCTGATCTG 2831
QY 2821 cctctccaaacttctgcgcgaagcgcccttcagccccctgcgcggagagagctctcc 2880
DB 2832 CCGTCCAACTTCCTTCCGCGCAAGCGGAGCGCCTTCAGCCCTTCGCGGAGAGTCTCC 2891
QY 2881 cgaagcagcggagccttcccgccatggtgagctcgcagcgtccagcgtgagcagcggcc 2940
DB 2892 CGAGAGCGCGGACGCTTCCGCGCCATGTTGGAGCTCCGAGGCTGGATTCGGAGCGCGCC 2951
QY 2941 ggggagcagcaggggtcctcttcgcgcggttctcgaagaccgagggcgagggcgaggg 3000
DB 2952 GGGAGCAGCGACGAGCTCTTTCGCGCGTCTCTCGAAGACCGAGCGGCGGAGCGCG 3011
QY 3001 ggcctctccagaccagaagctgaggaacctgtggtgagccgctgacatggaagatct 3060
DB 3012 GGCCTTCAGACCAAGAAAGCTTGAGAGCTTGGGTGAGCCCGCTGACCATGGAAAGATCT 3071
QY 3061 tgtctgtacagcttccaggtgagggaggtgaggttctctggcttcccgaagatgcat 3120
DB 3072 TGTCTGTACAGCTTCCAGTGGCCAGAGGATGGAGTTCCTTGGCTTCCGGAAGTGCAT 3131
QY 3121 ccacagagacctggtgctcggaacatctcgtgtcggaagcgagcgtggtgaagatctg 3180
DB 3132 CCACAGAGACCTTGGCTGCTCGGAACATTCCTGTGTGGAAAGCGAGCTGTTGAAGATCTG 3191
QY 3181 tgacttggccttgcggggagcatctacaagaacctgactacgtccgcaagggcagtgctc 3240
DB 3192 TGACTTTGGCTTGGCCGGGACATCTACAAAGACCTGACTACGTCCCGCAAGGCGAGTGC 3251
QY 3241 ccggctgccccctgaagtgagtgggccccctgaaagcatcttctgacaaggtgtacacacgca 3300


```

Db 3252 CCGGCTGCCCTGAAGTGGATGGCCCTGAAAGCATCTTCGACAAGGTGTACACCACGCA 3311
Qy 3301 gagtgcgtggtgctcttggagtgctctctctggagatctctctctctggggcctcccc 3360
Db 3312 GAGTGAGTGTGGTCTTTGGGGTGTCTCTGGAGATCTTCTCTCTGGGGCTTCCCC 3371
Qy 3361 gtacctggggtgcagatcaatgagaggttctgccagcgctgagagacgagcacaaggat 3420
Db 3372 GTACCTGGGTGCACATCAATGAGGAGTCTGCCAGCGCTGAGAGACGCGCACAAAGAT 3431
Qy 3421 gaggccccggaactggccactcccccacacacacacacacacacacacacacacacac 3480
Db 3432 GAGGGCCCCGAGCTGGCCACTCCCGCCATAGCCGCATCATGCTGAACGTGTGTGGTCCG 3491
Qy 3481 agccccaaagcgagacctgcattctcgagctggtggtgagatcctggggagctgctcca 3540
Db 3492 AGACCCCAAGCGAGACCTGCTCTCGGAGCTGTGGAGATCTTGGGGACTGCTCTCCA 3551
Qy 3541 gggcaggggctgcaagaaagagagaggtctgcatggtggcccgccgagctctcagagctc 3600
Db 3552 GGGCAGGGGCTGCAAGAGGAAGAGGAGTCTGCATGGCCCGCCGAGCTCTCAGAGCTC 3611
Qy 3601 agaagagggcagcttctcaggtgtccaccatggccctacacatgcgccaggtgagcgc 3660
Db 3612 AGAAGAGGGCAGCTTCTCGAGGTGTCCACCATGGCCCTTACACATGCCCGAGGCTGACGC 3671
Qy 3661 tgaggacagcccgccagcctgcagcgccacacagcctggcgccaggtattacaactgggt 3720
Db 3672 TGAGGACAGCCCGCAAGCTGCGAGCGCCACAGCCTGGCCGCCAGGTATTACAACTGGGT 3731
Qy 3721 gtccttccgggtgctggccagaggggctgagagcccggttctccagatgaagac 3780
Db 3732 GTCTTTTCCGGGTGCTGGCCAGAGGGCTGAGAGCGGTGAGACCGGTGCTTCCAGGATGAAG 3791
Qy 3781 attgaggaattcccatgaccccaacagacctacaagctctgtgacaaccaacaaga 3840
Db 3792 ATTGAGGAATTCCTGACCCCAACGACCTTACAAGGCTCTGTGGACAACACAGACAGA 3851
Qy 3841 cagtggagtggtgctggcctcgagaggtttgagcagatagagagcagcagcagacaga 3900
Db 3852 CAGTGGGATGCTGGCTCGAGGAGTTTGAGCAGATAGAGAGCAGCAGCATAGACAAGA 3911
Qy 3901 aagcggttcagct 3914
Db 3912 AAGCGGCTTCAGT 3925

```

```

RESULT 8
PCT-US95-04228-45
; Sequence 45, Application PC/TUS9504228
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04228
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/222616
; FILING DATE: 04-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wendy M. Lee
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 821P3PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9108 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US95-04228-45

```

```

Query Match 95.2%; Score 3912.4; DB 4; Length 9108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3913; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ccaacgcagcggcgagatgacgagggcgccgctgtgcctgcagctgctctg 60
Db 944 CCACGCGAGCGCGCGGAGATGCAGCGGGCGCGCGCTGCGCTGCGACTGTGCTCTG 1003
Qy 61 cctggagctcctggcggcctggtgagtggtactcctcagaccccccagacctgaacat 120
Db 1004 CCTGGGACTCTCTGGAGCGCTTGGTGGTGGTACTTCCATGATCCCCCGACCTTCAACAT 1063
Qy 121 cagcagagagtcacacgctcactgcacacccggtagacgctgtccatcctcaggggaca 180
Db 1064 CAGCGAGGAGTCCACGCTATCGACACCGGTGACAGCTGTCCATCTCTCGAGGGGACA 1123
Qy 181 gcaacccctcagtggtggtggcgagctcagagcgccagcgccagcgagacaga 240
Db 1124 GCACCCCTCTGAGTGGGCTTGGCCAGGAGCTCAGAGGGCGCGCAGCCGCGGAGACAAGA 1183
Qy 241 cagcagagacacgggggtggtgctgagactgagagggcagacgagcccgccctactgcaa 300
Db 1184 CAGCGAGGACACGGGGTGGTGGGAGACTGCGAGGGCACAGACGCGCCCTACTGCAA 1243
Qy 301 ggtgtgctgctgacagaggtacatgccaaacacacagcagcagctacgtctgctacaaa 360
Db 1244 GGTGTGTGCTGCGACGAGGTACATGCCAACACACAGCAGCTACGTCTGCTACTACAA 1303
Qy 361 gtacatcagggcacgcatcgaggggcaccacggccgagctcctcagctgttcgtgagaga 420
Db 1304 GTACATCAAGGCACCATCGAGGGCACACGCGCCGCGAGCTCTTCTAGTGTCTGTCAGAGA 1363
Qy 421 ctttgagcagccattcacaacacgctgacacgcttcttgggtcaacaggaagcagcact 480
Db 1364 CTTTGAGCAGCCATTTCATCAACAAGCTCTTGGTCAACAGAGAGAGAGCCCAT 1423
Qy 481 gtaggtgctgctggtgctgctcctcagcagcctcctcagcagcagcagcagcagc 540
Db 1424 GTGGGTGCGCTGTCTGGGTGCTCCTCCCGGCTCAATGTACGCTGCGCTCGCAAGCTTC 1483
Qy 541 ggtgctggtggcagcggcgagggaggtggtggtggatgacggcggggagctgctgttc 600
Db 1484 GGTGCTGTGGCCAGACGGCAGAGGTGGTGGGATGATGACCGGGGGGAGCATGCTGCTGTGTC 1543
Qy 601 cagcgcaactgctgacgagtgccctgtacctgagtgagagaccactgggagacacaga 660
Db 1544 CAGCGCACTGCTGCGACGATGGCCCTGTACTGCGAGTGGGAGACCACTGGGGAGACACAGA 1603
Qy 661 cttctcttcccaaccccttctcctggtgtagcatcacaggaacagagctctatgacatccagct 720

```

Db 1604 CTTCTTTCACCCCTCTCTGTGCACATCACAGCAACGAGCTCTATGACATCCAGCT 1663
QY 721 qtlgcccaggaagtcgtgagtcgtgtagggagagtcgttctgctgaactgcacct 780
Db 1664 GTTGCCAGGAAGTCTGCTGGAGCTGCTGGTAGGGAGAAGCTGGTCTGAATGCACTG 1723
QY 781 gtgggtgagtttaactcagggtgacacctttgactgggactaccaggaagacagga 840
Db 1724 GTGGGTGAGTTAACTCAGGTGTACCTTTGACTGGGACTACCCAGGGAAGCAGGACA 1783
QY 841 qcgggglaagtgggtcccgagcgagcgtccccagagagacccacacagaaactctcagact 900
Db 1784 GCGGGGTAAGTGGGTGCCGAGCGAGCGCTCCAGCAGACCCACACAGAACTCTCCAGCAT 1843
QY 901 cctgaccaaccacacacgctcagccagcagcagcctgggctcgtatgctgcaaggccaacaa 960
Db 1844 CCTGACCATCCACACGTCAGCCAGCAGCAGCAGCTGGGCTGCTATGTGTGCAAGGCCAACAA 1903
QY 961 cggcatccagcagattcgggagagaccagaggtcatctgctgcatgaaatcccccttcacag 1020
Db 1904 CGCATCTCAGCGATTTTCGGGAGAGCAGCAGGAGGTCTATTGTGCATGAAATCCCTTTCATCAG 1963
QY 1021 cgtcagtgctcaaaagacccatcctggagccacgagcagcagagcagcagcagcagcagc 1080
Db 1964 CGTGAGTGGCTCAAGGACCCATCTCTGGAGGCCAGCGCAGGAGCAGCAGCTGGTGAAGCT 2023
QY 1081 gcccgtgaagctggcagcgttacccccgcgcgcaggttccagtggttacaagagtggaagggc 1140
Db 2024 GCGCGTGAAGCTGGCAGCGTACCCCGCCCGAGTGTCCAGTGGTACAAAGGATGGAAAGGC 2083
QY 1141 actgtccggcgcacagtcacatgccctgctgtcgaagaggtgacagaagccagac 1200
Db 2084 ACTGTCCGGCGCCAGTCCACATGCCCTGGTGTCTCAAGGAGGTGACAGAGGCCAGCAC 2143
QY 1201 agcacctacacccctgcgcctgtgaaactccgctgctgagcctgagcagcagcagcagc 1260
Db 2144 AGGCACCTACACCTCGCCCTGTGGAACCTCCGCTGCTGGCCCTGAGGCGCAACATCAGCCT 2203
QY 1261 ggaagtcgttgtaagtgtcccccacagatacatagagaaggagggcctctccccagact 1320
Db 2204 GAGCTGTGGTGAATGTGCCCGCCAGATACATGAGAGGAGGCGCTCTCTCCCGCAGCAT 2263
QY 1321 ctactcgtctcagcgcgcagcgcctcactcgtcagcgcctcaggggtgcccctgcctct 1380
Db 2264 CTACTCGCGTCACAGCGCGCAGGCGCTCACCCTGACCGGCTACGGGGTCCCTCGCTCT 2323
QY 1381 cagcatccagtggcactggcgccctggacccctgcaagatgtttgcccagcgtagct 1440
Db 2324 CAGCATCCAGTGGCACTGGCGCCCTGGACACCTTGCAAGATGTTTGCCCAAGCTAGTCT 2383
QY 1441 ccggcgggcagcagcaagaactcatgcccagtgccgtgactggagggcggtgacacac 1500
Db 2384 CCGCGCGCGCAGCAAGACTCTATGCGACAGTGGCGTACTGGAGGCGGTGACCCAC 2443
QY 1501 gcsagatgcccgtgaaccccatcagagcctgacacctgacccagcttctgaggggaaa 1560
Db 2444 GCAGGATGCCGTGAACCCCATCGAGAGCTTGACACCTTGACCCGAGTGTGTGCGAGGAAA 2503
QY 1561 gaataagactgtgagcaagctggatccagaatgccacggtgtctgcatgtacaaagt 1620
Db 2504 GAATAAGACTGTGAGCAAGCTGTGATCCAGAAATGCCAAGCTGTCTGCTGCCATGTACAAAGTG 2563
QY 1621 tglggtctcaacaagtggtggcagatgagcgtcactctactctctctgtgacacacat 1680
Db 2564 TGTGGTCTCCAAAGTGGCGCAGATGAGCGGCTCATCTACTTCTATGTGACCCACCAT 2623
QY 1681 ccccgagcgtctcacatcgaatccaaagccatccgagagcgtactagaggggcagcagat 1740
Db 2624 CCCCAGCGGCTTCAACCTCAATCCAAAGCCATCCGAGGAGCTACTAGAGGCGCAGCCGGT 2683
QY 1741 gctccctgagctgccaagcagcagctacaagtacagacalcctgctggtaccgcctcaa 1800
|||||

Db 2684 GCTCTGAGCTGCCAAGCGCAGAGCTACAAGTACGAGCATCTGCGTGGTACCGCTCAA 2743
QY 1801 cctgtccacgctgcacgctgagcagcaggaacccgcttctgctcgaactcaagaacgtaca 1860
Db 2744 CTTGTCCACGCTGACAGATGCGCACCGGAACCCGCTTCTGCTCGACTGCAAGAAGTGA 2803
QY 1861 cctgttccacccctcttggccgcagcctggagaggtggcaccctggggcgccacgc 1920
Db 2804 TCTGTTGCCACCCCTCTTGGCGCCAGCCTGGAGGAGTGGACCTGGGGCGGCCACGC 2863
QY 1921 cagctcagcctgagatccccccgctgcgcccgcagcagcaggggccaactatgtgtgaga 1980
Db 2864 CAGCTCAGCTGAGATATCCCCCGCTGCGCCGAGCAGCAGGAGGCCACTATGTGTGGA 2923
QY 1981 agtcaagacccgctgagcctgacacgactgacacgactgcccacagagtcctgtcggtgagggc 2040
Db 2924 AGTGAAGACCGCGCAGCCATGACAGCAGTGCACAAAGTAACCTGTCTGCTGACGGC 2983
QY 2041 cctgaaagcccccggctcagcagaaacttgaccgacctcctggtaacgtgagcactc 2100
Db 2984 CTTGAAGCCCTCGCCTCAGCAGAACTTGACCGACCTCTGTGTGAACGTGAGCGACTC 3043
QY 2101 gctgagatgcaagtcttggtagccgagcagcagcagcagcagcagcagcagcagcagc 2160
Db 3044 CTTGAGATGCAAGTCTTGTGGCCGAGCGCAGCGCCCGCCAGCATGCTGTGTGTGTAAGA 3103
QY 2161 cagagggctgctggaggaagctggagtcgactggcagccttggcagcctggagcagc 2220
Db 3104 CGAGAGGCTGTGGAGGAAAGTCTGGAGTGGACTTGGCGACTCCAAACAGAGCTGAG 3163
QY 2221 catcagcagcgtgcagagagagatggggagcgtatctgtgcagcgtgtgcaacgcacaa 2280
Db 3164 CATCAGCGCGTGGCGAGAGGATGCGGGAGCGTATCTGTGACAGCGTGTGCAAGGCCAA 3223
QY 2281 gggctgctcactcctcccgagcgtggcgtggaaggtcccgaggaataagggcagcagc 2340
Db 3224 GGGCTGCGTCAACTCTCCGCGCAGCGTGGCGTGGAGGCTCCGAGGATAGGGCAGAT 3283
QY 2341 ggaagtcgtgactctgtgctacggcgctcagcgtctctctctctctctctctctctct 2400
Db 3284 GAGATCTGTGATCTCTGCGTACCGGCTCATCTGCTCTCTCTCTCTCTCTCTCTCTCT 3343
QY 2401 cctcatctctgtaacatgagagcggccacacacacacacacacacacacacacacacac 2460
Db 3344 CCTCATCTTGTAAATATGAGAGCGCGCCCGCCACGACACATCAAGAGCGGCTACTCTG 3403
QY 2461 catcatcagcagcccgggaggtgctctctgagagcaatgcaaatcctgtcctacga 2520
Db 3404 CATCATATGACCCCGGGAGTGGCTCTGGAGGAGCAATGCGAATACCTGTCTCTACGA 3463
QY 2521 tgcagcagctgggaattcccccgagagcgtgcacctggggagagtgctcggtacgg 2580
Db 3464 TCCAGCCAGTGGGAAATTTCCCGCAGAGCGGCTGCACCTGGGGAGTGTCTCGGCTACGG 3523
QY 2581 cgccttcgggaaggtggtagaagcctccgcttccgcatccacaggaagcaagcagctgga 2640
Db 3524 CCGCTTCGGGAAGTGGTGGAGGCGCTCCGCTTTCGGCATCCACAGGCGCAGCGCTGGA 3583
QY 2641 caccgtggcctgaaaaatgctgaaagagggcgccacgcccagcagcagcagcagcagcagc 2700
Db 3584 CACGCTGGCCGTGAAATGCTGAAAGAGGGCGCCACCGCCAGCAGCAGCAGCAGCAGCAG 3643
QY 2701 gtcggagctcagatcctcatcacaatcagcaacacacacacacacacacacacacacac 2760
Db 3644 GTCGGAGCTCAAGATCTCTTTCATTCATTCGGAACACCTCAACCTGGTGTCAACCTCTCTGG 3703
QY 2761 ggcgtgacacacagcgcaggg 2820
Db 3704 GCGGTGACCAAGCGCAGGG 3763
QY 2821 cctctcaactctctgc 2880
Db 3764 CTTCTCAACTTCTGCGCGCAAGCGGGAGCGCTTACGCCCTTACGCCCTTGGCGGAGAGTCTCC 3823

QY 2881 cagagcagcgcgacgtcttcgcgcgcacatggtgagctcgcacaggtggtatcgagagcgcc 2940
DB 3824 CGAGCAGCGGACGCTTCGCGGCCATGCTGAGCTCGCCAGGTGATCGAGCGGCC 3883
QY 2941 gggagcagcgcaggggtctcttcgcgcgcgttctcgaagaccgagcgagcgagcg 3000
DB 3884 GGGAGCAGCGACGGGTCTCTTCGCGGCTTCTCGAAGACCGAGGCGGAGCGAGCG 3943
QY 3001 ggccttcagcagcagagcgtgagagcctgtggtgagccgctgacatggaagatct 3060
DB 3944 GGCCTTCAGACCAAGAGCTGAGGACCTGTGGCTGAGCCGCTGACCATGGAAGATCT 4003
QY 3061 tctgtctacagcttcaggtggccagagggatggagttctctgcttcccaaaagtgcat 3120
DB 4004 TGTCTGCTACAGCTTCAGGTGCCAGAGGGATGGAGTTCCTGGCTTCGCCAAAGTGAT 4063
QY 3121 ccacagagacctggctcgcgaacatctcgtcgcgaagacgagcgtggaagatctg 3180
DB 4064 CCACAGAGACCTGGCTGCTCGGAACATCTGCTGCGAAGACGAGCTGGTGAAGATCTG 4123
QY 3181 tgacttggccttgcgcggagacatctacaaagaccctgactacgtcccgcaaggcgagtcg 3240
DB 4124 TGACTTTGGCCTTGGCCGGGACATCTACAAAGACCTGACTACGTCCGCAAGGCGAGTGC 4183
QY 3241 ccggctgccccgaagtgaatggagtcgccccgaagacatcttcgacaagggtgtacaccagca 3300
DB 4184 CCGGCTGCCCTGAAGTGGATGGCCCTGAAAGCATCTTCGACAAAGGTGTACACCACGCA 4243
QY 3301 gagtgcgtggtctcttgggggtctctctggagatctctctctggtggggtctccc 3360
DB 4244 GAGTGACGTGGTCTCTTTGGGGTCTCTCTGGAGATCTTCTCTCTGGGGGCTTCCC 4303
QY 3361 gtacctgggtgtagatcaatgagagttctgcccagcgctgagagcggcacaaagat 3420
DB 4304 GTACCTGGGGTGAGATCAATGAGAGTTCTGCCAGCGGCTGAGAGACGGCACAAGAT 4363
QY 3421 gaagccccgagctggcctccgcacatccgcacatccgcacatctgactgactgctcgg 3480
DB 4364 GAGGGCCCGAGCTGGCCACCTCCGCCATACGCCGCATGCTGTAAGTGTGCTGGTTCGG 4423
QY 3481 agaccccaaggcgagacctgcatcttcgagctggtgagatccctgggggacctgctcca 3540
DB 4424 AGACCCCAAGGCGAGACCTGCTATCTCGAGCTGTGGAGATCTCTGGGGACCTGCTCCA 4483
QY 3541 gggcagggcctcgaagaggaagagaggtctgcatgccccgcgcagcagctctcagagtc 3600
DB 4484 GGCAGGGGCTGCAAGAGGAAGAGAGGTCTGCAATGGCCCCGCGCAGCTCTCAGAGCTC 4543
QY 3601 agaagagggcagcttctcgcaggtgtccaccatggccctacacatcgcccaggtgagcg 3660
DB 4544 AGAAGAGGCGAGCTTCTCGAGGTGTCCACCATGGCCCTACACATCGCCCGAGCTGACGC 4603
QY 3661 tgaggacagccgcgaagcctcagcgccacagcctggcgccaggtattacaactgggt 3720
DB 4604 TGAGGACAGCCCGCAAGCGCTGAGCGCCACAGCCTGGCGCGCAGGTATTACAAGTGGGT 4663
QY 3721 gtccttccgggtgctgcccagaggggtgagaccgtggttctccagagtaagac 3780
DB 4664 GTCTTTCCGGGTGCTGGCCAGAGGGGCTGAGACCCGTGGTTCCTCCAGGATGAAGAC 4723
QY 3781 atttgaggaaattcccatgacccccacacacacacacacacacacacacacacacacacac 3840
DB 4724 ATTTGAGGAATTCCTCCATGACCCCAACGACCTTACAAAGCTCTGTGGACAAACAGACAGA 4783
QY 3841 caatggagatggtgctgagcctcgagaggtttgagcagatagagcagcagcagcagcagc 3900
DB 4784 CAGTGGGATGGTGTGGCTCGAGGAGTTTGAGCAGATAGAGCAGGCGATAGACAAGA 4843
QY 3901 aagcggcttcagct 3914
DB 4844 AAGCGGCTTCAGCT 4857

RESULT 9
US-08-222-616-17/c
: Sequence 17, Application US/08222616
: Patent No. 5635177
: GENERAL INFORMATION:
: APPLICANT: Bennett, Brian D.
: APPLICANT: Goeddel, David
: APPLICANT: Lee, James M.
: APPLICANT: Matthews, William
: APPLICANT: Tsai, Siao Ping
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
: TITLE OF INVENTION: ANTIBODIES
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION NUMBER: US/08/222.616
: FILING DATE: 4-APR-1994
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/00586
: FILING DATE: 22-JAN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/826935
: FILING DATE: 22-JAN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Lee, Wendy M.
: REGISTRATION NUMBER:
: REFERENCE/DOCKET NUMBER: 821P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 17:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6827 bases
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-222-616-17

Query Match 25.3% Score 1040.2; DB 1; Length 6827;
Best Local Similarity 98.3%; Pred. No. 1.4e-203;
Matches 1051; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 2846 cggagcgccttcagccctcgcgagagagtgctcccgagcagcgagcgttcgcgcgc 2905
DB 2947 CGAGCAGCTTCGTGTCATCCCGCAGGAGAGTCTCCCGAGCAGCGGACGCTTCGCGCC 2888
QY 2906 atggtggagctcgcagcagctgagtcgagcgcgcgagcgagcgagcgagcgagcgagcg 2965
DB 2887 ATGTTGGAGCTCGCCAGGCTGATCGAGCGCGCGGAGCGAGCGAGGCTCTCTTC 2828
QY 2966 gcgcgggttctcgaagaccgagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 3025
DB 2827 GCGCGGTTCTCCAGACCGAGGCGGAGCGAGGCGGCGGCTTCTCCACAGCAAGAGCTGAG 2768
QY 3026 gacctgtgagcgcgcgtaccatgaaagatcttctgctcagctcagctccagctgacc 3085
DB 2767 GACCTGTGCTGAGCCCGCTGACCATGGGAAGATCTTGTCTGTACAGCTTCCAGGTGCC 2708


```

: Sequence 5, Application US/07977451
: Patent No. 5270458
: GENERAL INFORMATION:
: APPLICANT: Lemischka, Ihor R.
: TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
: NUMBER OF INVENTIONS: 1
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ImClone Systems Incorporated
: STREET: 180 Varick Street
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10014
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/977,451
: FILING DATE: 19921119
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US UNASSIGNED
: FILING DATE: 12-NOV-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/906,397
: FILING DATE: 26-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US PCT/US92/05401
: FILING DATE: 26-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: TW 81102961
: FILING DATE: 15-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US PCT/US92/02750
: FILING DATE: 02-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/813,593
: FILING DATE: 24-DEC-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/793,065
: FILING DATE: 15-NOV-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/728,913
: FILING DATE: 28-JUN-1991
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/679,666
: FILING DATE: 02-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Feit, Irving N.
: REGISTRATION NUMBER: 28,601
: REFERENCE/DOCKET NUMBER: LEM-3-7P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-645-1405
: TELEFAX: 212-645-2054
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5406 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 208..4311
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 265..4308

```

:	FEATURE:
:	NAME/KEY: sig_peptide
:	LOCATION: 208..264
:US-07-977-451-5	
Query Match 17.3%; Score 713; DB 1; Length 5406;	
Best Local Similarity 54.5%; Pred. No. 7.4e-137;	
Matches 1711; Conservative 0; Mismatches 1345; Indels 81; Gaps	
QY	464 acagagaagcgcgatgtgggtgcccctgtcgttgggtccatcccgccccaatgtcaacg 523
DB	
QY	524 ctgcgc-----tcgcaaaagtccgtgtcgttgccagacaggcagggaggtgggtgtaggat 577
DB	
QY	578 gaccggcggggcatgctcgtgccagcccaactgctgcacgatgccttgtacctgcagtcg 637
DB	
QY	637 agcgagataggcttttactctccccagtttacatgatcacactatgccggcatggcttctctgt 813
DB	
QY	638 gagcacacctggggagaccaggaactcccttccaacccttccctgggtgcacatcacaggc 697
DB	
QY	698 aacgagctctatgacatccagctgtgtgccaggaagtcgctggagctgtcgttggtagggag 757
DB	
QY	758 aagctggctgaactgcaccgtgtgggtgagtttaactcaggtgtcacctttgactgg 817
DB	
QY	817 aaacttgccttaaattgtacacgagaaacagagctcaatgtggggcttgatttcacctgg 993
DB	
QY	818 gaclaccccgaggaaacagcaggc-----agagcggggtaagtgggtgcccgagcagcgt 869
DB	
QY	869 cactgtccacacttcaaagcttcataagaaagattgttaaaccgggatgtgaaaccccttt 1053
DB	
QY	870 ccagcagaccacacagaactctcc-agcatcctlgaccatccacaagtlcagccagcac 928
DB	
QY	928 gacctgggtcgtatgtgtgcgaaggccaacacggcatccagcatttcgggagagcacc 988
DB	
QY	988 gaggctatgtcatgaaaaatcccttcacagctcaggtgcagttgcctaaggaccatccctg 1048
DB	
QY	1048 ttgttccgagtttcacaaaagccttttatgtcgttgcgtatgggatgaaattctttgggtg 1233
DB	
QY	1049 gaggccacggcagagcagcgtgttgaagctgcctgtgaagctggcagcgtcacccccc 1108
DB	
QY	1108 caagacca--cagtgggcagtcgaagtccgaattccctgtgtgaagtatctcagttaccac 1290
DB	
QY	1109 ccgaggttccagtggtataaggatgaaagcactgtccggggccacsa-----gt 1159
DB	
QY	1291 cctgatcaaatgggtacaaaattggaaagcccaattgagttccaaactacaaatgattgtt 1350
DB	
QY	1160 ccacatlqccctggctcaaggaggtgcagagggccagcacagcagcactacacccctcgcc 1219
DB	
QY	1351 ggccgatgaactcaccatcatcgaagtgcattcaagagatgcagaaactacacgggtcattc 1410
DB	
QY	1220 ctglgaaactccgctgctggtccaggcgcacatcagccttgagctgggtggatgtg 1279
DB	
QY	1411 ctaccacaccccatTTCAAATGGAGAAACAGAGCCACATGGTCTCTCTGTTGTGAATGTC 1470
DB	
QY	1280 cccccccagatcatgagaaggagcctcttccccca---gcactcactcgctcacagc 1336
DB	
QY	1471 CCACCACCATCGGTGAGNAAGCCTTGATCTCGGCTATGGGATTCTTACCAGTATGGGACC 1530
DB	
QY	1337 cgccaggccctcacctgcacggcctaacgggtgccccctgocctctcagcatcccaagtggcac 1396
DB	
QY	1531 ATGCACACATTCGACATGCACAGCTTACGCCAACCCCTCCCTGCAACCACTACAGTGGTAC 1590
DB	

Db 1747 ACCTGCTATCAAGCTGCCAACGTGTTCAGCGTTTGTACAAATGTGAAGCCATCAACAAA 1806
Qy 1637 gtggccaggatagcggtctatctactctatctatgtgaacacacaccccgagcgtctcacc 1696
Db 1807 GCGGACGAGGAGAGAGGTCTCTCTTCCATGTGATCAGGGGTCTCTGA-----A 1857
Qy 1697 atcgaaatccaaagccatccagagagctactagagggccagcgggtgtctctctgagctgccc 1756
Db 1858 ATACTGTGCAACCTGCTGCCACCACTGACGAGAGAGTGTCTCTCTGTTGTGCACT 1917
Qy 1757 gccacagctacagtagcatctgcgtgtagtaccgacctcaacctgtccacgtgacac 1816
Db 1918 GCACAGAGAAATACGTTTGAGAACTCAGGTGTGTAAGAGCTTGCTTCACAGGCAAGTTG 1977
Qy 1817 gatcgcaagggaaacggctctctgctgactcaagaacgtgcatctgttgcgccacct 1876
Db 1978 GTCCACATGGCGCAATCACTACACCAAGTTTGCAGAACTTGGATGCTCTTTTGGAAACTG 2037
Qy 1877 ctggcccgccagcctggaggggtggcaccttggggcgccacccacgcctcagcctgagt 1936
Db 2038 AATGGCACCA-----TGTTTCTACAGCACAAATCACATCTTGTATTGTGGCA 2085
Qy 1937 atcccccgctgcgcccgcagcagggccactatgtgtgcaagtgcgaagccggcg 1996
Db 2086 TTTGAGATGCTCTCTGCAAGGACCAAGGCGACTATGTTTGTCTGCTCAAGATAAGAAG 2145
Qy 1997 agccatgacaagcactgccacaagaagtacctgtcggtgcagccctggagccctcgg 2056
Db 2146 ACCAAGAAAGACATTTGCTGTGCTCAACACGCTCATCTCTAGAGCCATGGCACCCCATG 2205
Qy 2057 ctacgcagaaacttgaccgacctctctggtgaacgtgagcactgcgtgagatgcagtlgc 2116
Db 2206 ATACCGGAAATCTGGAGATTCAGAACACCAACCACTTGGCGAGACCAATGGAATGACTTGC 2265
Qy 2117 ttggtggcgagcgcaagcccgagcatcgtgtgttacaagaagagagcgtgctggag 2176
Db 2266 CCAGCATCTGGAATTCCTACCCACACATTTACATGTTGTTCAAGACACACAGACCTTGTA 2325
Qy 2177 gaaagtgtgagtcgacttggcgactcccaaccagaagctgagcactccagcgctgcgc 2236
Db 2326 GAAGATTACGACCATTTACTGAGAGATGGGAACCGGAACCTGACTATCCGACGGGTGAG 2385
Qy 2237 gagagagtgagcagcgtactgtgagcggtggaacccagcagcagcgtgcgtcaactcc 2296
Db 2386 AAGGAGATGAGGCGCTTACACCTGCGAGCGCTTGCATGCTTGGCTGTGCAAGAGCG 2445
Qy 2297 tccgcagcgtggcggtggaggtccgaggaagaagggcagcagcagcagcagcagcagc 2356
Db 2446 GAGACGCTTTCATATAGAGAGTGGCCAGGAAAGACCACTTGGAAATGATATATCTC 2505
Qy 2357 gtccgtaccggtgcatcgtgtcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2416
Db 2506 GTCCGCACTGCACTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2565
Qy 2417 atgagagcgccgcccacgacagcatcaagcagcagcagcagcagcagcagcagcagc 2476
Db 2566 GTTAACCGGCAATGAGAGGGAATGNAGACAGGCTACTTGTCTATTGTCTATGATGCA 2625
Qy 2477 ggggaggtgctctggaggaagcaatccctgtcctacagcagcagcagcagcagcagcagc 2536
Db 2626 GATGAATTCCTTGGATGAGCGTGTGAGCGCTTGGCTTATGATGCGCAGCAAGTGGAA 2685
Qy 2537 tcccccgagagcggtcagcctggggagagtgctcggtcagcggcgtcgggaggtg 2596
Db 2686 TTCCCGAGGAGCGGCTGAAACTAGGAAACCTCTTGGCCCGGTGCTTGGCCCAAGT 2745
Qy 2597 gtggaagcctcgtctcgcatcccaagggcagcagcagcagcagcagcagcagcagcagc 2656
Db 2746 ATTGAGGAGAGCGCTTTTGGAAATGACAAAGACAGCGACTTGCAGAAACAGTAGCCG 2805
Qy 2657 atgctgaaagagggcgccagcggcagcagcagcagcagcagcagcagcagcagcagcagc 2716
Db 2806 ATGTTGAAAGAGGAGCAACACAGCAGCGCATCGAGCCCTCATGCTGTGAACCTCAAGATC 2865

Qy 2717 ctcaatcacatcgcaacacccctcaacgltggtcaacctctcggggcggtgacccaagcgg 2776
Db 2866 CTATCCACATTTGGTCACCATCTCAATGTGTGTGAACCTCTAGGCGCTGACCAAGCGC 2925
Qy 2777 caagggccctcatggtgatcgtgaggttctgcaagtagcagcaacctctcaactctcgt 2836
Db 2926 GGAGGGCCTCTCATGTTGATTTGGAATTTCTGAAATTTGGAACCTATCAACTTACTTTA 2985
Qy 2837 cgcgcgaagcgagcgcttcagccccctgcgcggaagaagctccccgagcagcgcgagcgc 2896
Db 2986 CGGGGCAAGAGAAATGAATTTGTTCCCTATAGACAAAGGGCGACGCTTCGCCAGGCG 3045
Qy 2897 tlcgcggccatggtgagcgtccagggctggatcgagggcgggcgggagcgagcgagcgc 2956
Db 3046 AAGGACTACGTTGGGAGCTCTCCGTGGATCTGAAAGACGCTTGACAGCATCACACAGC 3105
Qy 2957 gtcctctcgcgcggttctcgaagaccgagggcgagcgagcgaggggcttctccagaccaa 3016
Db 3106 AGCCAGAGCTCTGCCAGCTCAGGCTTTGTTGAGGAGAAATCGCTCAGTGATGTAGAGAA 3165
Qy 3017 gaag-----ctgagagacctgtgctgagcccgctgacctggaagatcttctgctgctac 3070
Db 3166 GAAGAAGCTTCTGAAGAACTGTACAGGACTTCTCTGACCTTGAGCATCTCATCTGTTAC 3225
Qy 3071 agcttccaggtggccagagggatggagttcctgggttcccgaaagtgcacatccacagagac 3130
Db 3226 AGCTTCAAGTGGCTAAGGCACTGAGTCTTGGCATCAAGGAAGTGTATCCACAGGGAC 3285
Qy 3131 ctgctgtcgcgaacattctgtcgaagcgagcgagcgagcgagcgagcgagcgagcgagcg 3190
Db 3286 CTGGAGCAGCAAACTTCTCTATCGAGAAAGATGTGTTAAGATCTGTGACTTCCGCGC 3345
Qy 3191 ctggccgggacatcacaaagacctgactacgtccgcaagggcgagtgcccggtgcgc 3250
Db 3346 TTGGCCCGGACATTTATAAGACCCGGATTTATGTGAGAAAGAGAGATGCCGACTCCCT 3405
Qy 3251 ctgaaatgagtgccctgaaagcatcttcgaaaggtgtacacacgacagagtagcgtg 3310
Db 3406 TTGAAGTGTATGCCCGGAAACCAATTTTGACAGAGTATACACAAATTCAGAGCGATGTG 3465
Qy 3311 tgcctcttgggtgctctctggtgagatctctctggtggcgctcccgctaccctggg 3370
Db 3466 TGGCTTTCGCTGTGCTCTGGAATATTTTCTTAGTGGTCCCTCCCATACCTGGG 3525
Qy 3371 gtcagatcaatgagaggttctcccgcgctgagagcgagcagcaaggtgagggcgcccg 3430
Db 3526 GTCAAGATTGATGAAGAAATTTGTAGAGATTGAAAGAGGAACTAGAAATGCGGGCTCT 3585
Qy 3431 ggcgtggccactcccgccataccgcatcatgctgaactgctggcgagagaccccaag 3490
Db 3586 GACTACTACTACCAGAAATGTACCAGCATCTGTGACTGCTGGATGAGGACCCCAAC 3645
Qy 3491 gcgagacgtcattctcggagcgtggtgagatcctcgggggagcctgctccagggcagggcg 3550
Db 3646 CAGAGACCCCTCGTTTTCAGAGTTGGTGGAGCATTTGGGAAACCTCTGCAAGCAAAATGCG 3705
Qy 3551 ctgcaagaggaagagga 3567
Db 3706 CAGCAGGATGGCAAGA 3722

RESULT 15

US-08-252-517-5

; Sequence 5, Application US/08252517

; Patent No. 5548065

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ImClone Systems Incorporated

Qy	1517	cccatcgagagcctggacacctggaccagttgttgagggggaagaataagacctgtgagc	1576
Db	1687	AAGATCGAAGTCACCAAAACCAATATCCCTGATTGAAGGAAAAACAATACTGTAAGT	1746
Qy	1577	aagctggtgatccagaatgccaaagctgctcccatgtacaagtgctggtgctccaacaag	1636
Db	1747	ACCTGGTCATCCAACTGCCAACGTGTACGGTGTGTACAANTGTGAAGCCATCAACAA	1806
Qy	1637	gtggccagatgagcggctcatctacttctatgtgaccaccaLccccegcgcttcacc	1696
Db	1807	CGGGACGAGGAGAGGGGTCACTCTCCATGTGATCAGGGGTCTCTGA- - - - - A	1857
Qy	1697	atcgaatccaagccatccgagagagctactagagggccagcgggtgctctctgagctgccaa	1756
Db	1858	ATTACTGTGCAACTCTGCCCCAGCAACTGAGCAGGAGAGTGTGCTCTTTGTGCAC	1917
Qy	1757	gccgacagctLacaagtacagcatctgcgtggtaccgcctcaacctgtccacgctgcac	1816
Db	1918	GCAGACAGAAATACGTTTGAGAACCTCAGTGGTGTACAAGCTTGCTCACAGGCAACATCG	1977
Qy	1817	gatgcgaocgggaacccgcctctctctgactgcgaagaacgtgcactctgttcgcacccct	1876
Db	1978	GTCCACATGGCGGAATCACTCACACAGATTGTGCAAGAACTTGGATGCTCTTTGGAAC	2037
Qy	1877	ctggcccgacgctggagaggtggcaactgggcgcgcacgcctcagcctcagcttgagt	1936
Db	2038	AAATGGCACCA- - - - -TGTTTTCTAACACCAAAATGACATCTTGATTGTGGCA	2085
Qy	1937	atccccgcgtcgccgcgagcacagggccactatgctgcgaagtgcgaagccagcgcg	1996
Db	2086	TTTCAGAAATGCCCTCTCGCAGGACCAAGCGACTATGTTGCTCTGCTCAAGATAAAGA	2145
Qy	1997	agcatgacaagcactgccacaagaagtacctgctcggtgcagcgccctggaaagccctcgg	2056
Db	2146	ACCAAGAAAGACATTGCTGGTGTCAACAGCTCATCATCTTAGAGCGCATGGCACCCATG	2205
Qy	2057	ctcacgcagaacttgaccgacctctggtgaagctgagcgactcgctggagatgcagctgc	2116
Db	2206	ATCACCGGAAATCTCGAGAAATCAGACAACAACCATTTGCGAGAGCATTGAAGTGACTTGC	2265
Qy	2117	ttggtggcgaagcgacgcgccagcatcgtggttacaagaagcagagagcgtcgtggag	2176
Db	2266	CCAGCATCTGGAANTCCTACCCACACATTCATAGTTTCAAGAGACAAGACCCCTGGTA	2325
Qy	2177	gaaaagcttgagctgacttggcactccaaccagaagctgagcatccagcgctgcgc	2236
Db	2326	GAAGATTCAGGCATTGTACTGAGATGGGAACCGGAACCTGACTATCCCGCAGGTGAGG	2385
Qy	2237	gaggagatcgggagcgtatctgtgcagcgtgtgtcaacgccaggcgtcgctcaactcc	2296
Db	2386	AAGGAGATGGAGGCCTCTATACCTTGCCAGGCCTGCAATGCTCTTGCTGTGCAAGAGC	2445
Qy	2297	tcgcgcagcgtggcctggaaagcctccaggataaaggcagcatggagatcgtgactcctt	2356
Db	2446	GAGACGCTCTTCATATAGAAAGTCCCGAGGAAGACCAACTTGGAAAGTCATTATCTCTC	2505
Qy	2357	gtcggtagccgcgtcatcgtctctctgtgggtccctcctcctcctcctctctctgtaac	2416
Db	2506	GTGCGCACTCAGTGATGCCATGTCTTCTGGCTCCTCTTGTCTTGTCTCTCGTACGGACC	2565
Qy	2417	ataagagggccgcccacgcacatcaagcagggctacctgtccatcatcatgagccccc	2476
Db	2566	GTTAAGCGGCGCAATGAAGGGGAACGTGAAGCAGCGCTACTTGTCTATTGTATGGATCCA	2625
Qy	2477	ggggaggtgcctctggaggaacatgcgaaLacctgtcctacgatgccagccagtgaggaa	2536
Db	2626	GATCAATTGCCCTTGGATGAGCGCTGTGAACCTTGGCCTTATGATGCCAGCAACTGGGAA	2685
Qy	2537	ttccccgaagcgcgtgcacctggggagagtgctcggctacgcgcctctcgggaagtg	2596
Db	2686	TTCCCCAGGACCGGCTGAACATAGAAAAACCTCTTGCCCGCGGTGCTCTCGGCGCAAGT	2745
Qy	2597	gtggaagcctccgcttttcggcatccacaagggcagcagcatgtgacacgctggccgtgaaa	2656

[illegible]

flk-1. Mus muscul
Flkl receptor prot
Sequence of murine
Protein tyrosine-k
VEGF receptor KDR
Truncated KID sVEG
Human soluble vasc
Soluble truncated
Anti-angiogenic Fl
Human VEGFR-2 cata
Human VEGFR-2 cata
KDR protein sequen
Human VEGF recepto
Human VEGF recepto
Alpha type PDGF re
Platelet derived g
Truncated FLT sVEG
Soluble truncated
Human soluble vasc
Human VEGF recepto
Amino-terminal fra
Type B human plate
Platelet-derived g
Human platelet-der
Human sVEGF-RII.
Soluble truncated
Human soluble vasc
Bovine c-kit bk-1
Murine flk-2 recep
Murine flk-2. Mus
Murine flk-2. Mus
Flk2 receptor prot

13 2662 36.9 1367 13 R28041
14 2659 36.9 1367 16 R67817
15 2642 36.6 1367 15 R54046
16 1791 24.8 348 16 R85938
17 1476.5 20.5 848 18 W38240
18 1329.5 18.4 789 15 R62488
19 1329.5 18.4 789 20 W68007
20 1320 18.3 788 19 W47038
21 1315.5 18.2 805 20 Y06196
22 1284 17.8 366 21 Y87421
23 1256 17.4 316 21 Y87417
24 1252.5 17.4 764 19 W69679
25 1218 16.9 738 20 Y29639
26 1218 16.9 738 21 Y59302
27 1163.5 16.1 1089 11 R06910
28 1162.5 16.1 1089 11 R08267
29 1137.5 15.8 780 15 R62487
30 1137.5 15.8 780 19 W47039
31 1137.5 15.8 780 20 W68008
32 1092.5 15.2 758 18 W36138
33 1092.5 15.2 758 18 W13395
34 1085.5 15.1 1106 13 R26205
35 1084 15.0 1106 10 P90127
36 1084 15.0 1106 17 R99690
37 1051 14.6 1106 10 P90646
38 1033.5 14.3 664 15 R62486
39 1033.5 14.3 668 19 W47037
40 1033.5 14.3 668 20 W68006
41 1024 14.2 977 21 Y51322
42 954.5 13.2 992 14 R44994
43 954.5 13.2 992 14 R31375
44 954.5 13.2 992 14 R37502
45 954.5 13.2 992 16 R67815

ALIGNMENTS

RESULT 1
R85937
ID R85937 standard; Protein: 1298 AA.
XX
AC R85937:
XX AC
XX 14-FEB-1996 (first entry)
XX Protein tyrosine-kinase SAL-S1.
DE
XX Protein tyrosine-kinase; SAL-S1: cell growth; differentiation.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..24
FT /label= Sig_peptide
FT Domain 25..775
FT /label= Extracellular_domain
FT Domain 776..800
FT /label= Transmembrane_domain
FT Domain 801..1298
FT /label= Cytoplasmic_tyrosine_kinase_domain
XX W09527061-A1.
PN FLT4 receptor tyro
XX Human receptor tyr
XX Novel type III RTK
XX Human KDR protein.
XX Murine flk-1 recep
XX Human flk-1. Homo
XX Murine flk-1. Mus
XX Mouse foetal live
XX Murine foetal live
XX Murine flk-1 recep
XX Murine flk-1 prote

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM protein - protein search, using sw model
Run on: January 23, 2001, 03:45:12 ; Search time 73.6 Seconds
(without alignments)
633.236 Million cell updates/sec

Title: US-09-375-248-2
Perfect score: 7211
Sequence: 1 MORGALCLRLWLCGLLDG.....SEEDHCSPSARVTFITDINSY 1363
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 268485 seqs, 34193795 residues
Total number of hits satisfying chosen parameters: 268485
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6852	95.0	1298	16 R85937	Protein tyrosine-k
2	6852	95.0	1298	17 R90528	FLT4 receptor tyro
3	2753	38.2	1355	20 W80997	Human receptor tyr
4	2753	38.2	1356	13 R26999	Novel type III RTK
5	2753	38.2	1356	19 W59275	Human KDR protein.
6	2665	37.0	1367	14 R44996	Murine flk-1 recep
7	2665	37.0	1367	14 R41377	Human flk-1. Homo
8	2665	37.0	1367	14 R37504	Murine flk-1. Mus
9	2665	37.0	1367	16 R67537	Mouse foetal live
10	2665	37.0	1367	17 R97420	Murine foetal live
11	2665	37.0	1367	18 W19875	Murine flk-1 recep
12	2665	37.0	1367	20 Y08618	Murine flk-1 prote

Bennett ED, Goeddel D, Lee JM, Matthews W, Tsai SP;

PI Wood WI:
XX WPI: 1995-366160/47.
DR N-PSDB; T03090.
XX
XX Agonist antibodies which activate specific protein tyrosine
PT kinase(s) - also activate chimeric proteins of kinase extracellular
PT domain and Ig constant domain, useful for studying, and therapeutic
PT modulation of, cell growth and differentiation
XX
XX Disclosure; Fig 15A-F; 125pp; English.
PS
XX DNA probes based on protein tyrosine-kinase (PTK) sequences were used
CC to screen cDNA libraries to identify novel PTK genes. The SAL-S1
CC gene was isolated from several megakaryocytic cell line libraries.
CC The gene (see T03090) encodes a protein (R85937) showing homology to
CC FLT/PLK family PTKs. SAL-S1 can be used to target proteins in
CC drug design, as an (ant)agonist of other PTKs, to screen PTK
CC ligands or to modulate megakaryocyte/platelet adhesion.
XX
SQ Sequence 1298 AA:
Query Match 95.0%; Score 6852; DB 16; Length 1298;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MORGAAALCLRLWICGLLDGLVSGSYWTPPTLNITEESHVDTGDSLSISCRGHPLEWA 60
DB 1 mdrgaalclrlwlcglldglvsgsywtpptlniteeshvdtgdslsiscrgqhplewa 60
QY 61 WPGAQAPATGDKOSEDGVVRCDEGTDARPYCKVLLHEVHNDTGSYVCYKYIKARI 120
DB 61 wpgaapatagdksedgvvrcegdtdarpyckvllhevhandtgsyvcykyikari 120
QY 121 EGTAAASYVVRDEPQFINKPDTLLNKRDKAMVCLVSLPGLNLTLSOSSVLWPDG 180
DB 121 egttaasyvvrdepqfinkpdtllnkrdkamvclvslpglnltsossvlwpgd 180
QY 181 QEVWDDRRGMLVSPNLLHDALYLQCETTWGDQDFLSNPFLVHITGNELYDQLPKSL 240
DB 181 qevwddrrgmvlspnllhdalylqcettwgdqdflnspflvhitgnelydqlprksl 240
QY 241 ELLVGEKLVNCTVWAEVNSGVTFDWDYPGKAERGKWPERRSOOTHLSLTLTHNV 300
DB 241 ellvgeklvlnctvwaevnsgvtfwdypgkaergkwperroothlsstlthnv 300
QY 301 SOHDLGSYCKANNGIOREESTEVIVHENPFISVEWLKGPTEATAGDELVKLPVKLAA 360
DB 301 sohdlgsyckanngioreestevivhenpfisvewlkgpilateatagdelvklpvlkaa 360
QY 361 YPPPEFOWYKDGKALSGRSHALVULKEVTEASTGTYYTLALWNSAAGLRNTSLVNVN 420
DB 361 ypppefowkydgkalsgrshalvulkevteastgttytlalwnsaaglrntslvvnv 420
QY 421 PPQIHEKEASSPSIYSHRSQALCTAYGVPLPLSTOMHWRPWTCKMFAQSLRRROOQ 480
DB 421 ppqihekeasspsiyshrsqalctaygvplplstomhwrpwtckmfaqslrrrqoq 480
QY 481 DLMPQCDMRVATODAVNPISLDTWTEFEVGGKNTVSKLVIONANVSAMKCVVSNKV 540
DB 481 dlmpqcdmravtodaavnplsldtwtefevgkntvsklvionanvsamkcvvsnkv 540
QY 541 QODERLIYFVVTIPDGFTTESPSELEGGOPVLLSCQADSYKYEHLRWYRLNLTSLHD 600
DB 541 qoderliyfvvtipdgfttespseleeggopvllscqadskyehlrwyrlnltslhd 600
QY 601 AHGNPLLLDCKNHLFATPLAASLEEVAPGARHATLSLIPRVAPHEGHYCEVQDRRS 660
DB 601 ahgnpllldcknhlfatplaasleevapgarhatlsisiprvapeghyvevqdrsr 660
QY 661 HDKHCCHKYLSVQALEAPRLTONLTDLVNVSDSLEMOCLVAGAHAPSIVWYKDERLLEE 720
DB 661 hdkhchkylsvqaleaprltonltdlvnvdsdlemoclvaahapsivwykderlle 720

DB 661 hdkhchkylsvqaleaprltonltdlvnvdsdlemoclvaahapsivwykderlle 720
QY 721 KSGVDLADSNOKLSIORVREEDAGRYLCVNAKGCVNSSASVAVEGSEDKSGSHEIVLV 780
DB 721 ksgvdladsnoklsiorvreedagrylcsvnakgcvcnssasvavegsedkgsmeivlv 780
QY 781 GTGVIAVFWVLLLIIFCNMRRPAHADIKTGYLSIIMDGPVPLEEOCEVLSYDASQWEF 840
DB 781 gtgviafvfwvllliifcnmrrpahadiktgylsiimdgpvpleeocelysydasqewef 840
QY 841 PRERLHILGRVLGYCAFCKVVEASAFGIHKSSCDTAVKMKLEGATASEHRAIMSEIKIL 900
DB 841 prerlhilgrvlgysaigkveasafgihksscdtavkmlkegatasehramseikil 900
QY 901 THICNHLNVNMLIGACTKPOGPLMWIVVEFCYKGNLSNFRRAKRDFAFSPCAEKSPEORGRF 960
DB 901 thicnhlnvnmiligactkpgoplmwivvefcykgnlisnfrakrdafspcaekspeqgrf 960
QY 961 RAMVELARLDRRRPGSSDRVLFAFSEKTEGGARRASPDQEAEDLWLSPLTMEDLVCSYFQ 1020
DB 961 ramvelarlldrppgssdrvlfafsekteggarraspdqeaedlwspltmedlvcsyf 1020
QY 1021 VARGMEFLASRKCIIHRDLAARNILLESVDVVKICDFGLARDIYKDPDYVRKGSARLPLKW 1080
DB 1021 vargmeflasrkciihrdlaarnillessdvvkvicdfglardiykdpdyvrkgsaripkw 1080
QY 1081 MAPESIFDKVYVTTQSDVMSFGVLLWEIFSLGASPPYGVQIINEEFCQRLRDXTRMRAPELA 1140
DB 1081 mapesifdkvyvttqsdvmsfgvllweifslgasppgyqvqineefcqrldtrmrapela 1140
QY 1141 TPAIRRMILNCWGGOPKARPAFSELVEITLGLDGLQGRLOEEVEECMAPRSSOSSEGSFS 1200
DB 1141 tpairrimlncwggopkarpafselveitlglldglqgrloeeveecmaprssossegsfs 1200
QY 1201 QVSTMALHIAQADAEDSPSLORHSLAARYNWNVSFPGCLARGAETRGSSRMKTFFEEFPM 1260
DB 1201 qvstmalhiaqadaedspslorhslaarynwnvsfpgclargaetrgssrmktfeefpm 1260
QY 1261 TPTTYKSGVDNOTDSGMVLASEEFOIESRHRROESGF 1297
DB 1261 tpttyksgvdnotdsgmvlaseefoiesrhrroesgf 1297
RESULT 2
R90528
ID R90528 standard; Protein; 1298 AA.
XX
AC R90528;
XX
DT 08-APR-1996 (first entry)
XX
DE FLT4 receptor tyrosine kinase.
XX
KW FLT4: receptor tyrosine kinase; metastasis; lymphoma; lymphangioloma;
KW immunological disease; cancer; diagnosis; therapy;
KW monoclonal antibody.
XX
OS Homo sapiens.
XX
PN W09533772-A1.
XX
PD 14-DEC-1995.
XX
PF 09-JUN-1995; 95WO-F100337.
XX
PR 09-JUN-1994; 94US-0257754.
XX
PA (ALIT/) ALITALO K.
PA (KAIP/) KAIPAINEN A.
PA (KARN/) KARNANI P.
PA (KORH/) KORHONEN J.
PA (MATI/) MATTIKAINEN M.
PA (MUST/) MUSTONEN T.

PA (PAJU/) PAJUSOLA K.
 XX Alitao K, Kaipainen A, Karnani P, Korhonen J, Matikainen M;
 PI Mustonen T, Pajusola K;
 XX WPI: 1996-040189/04.
 DR N-PSDB: t12068.
 XX Antibodies against FLT4 receptor tyrosine kinase, hybridomas and
 PT nucleic acid probes - used to diagnose and treat e.g. metastatic
 PT cancers, involving alterations to lymphatic vessels.
 XX Disclosure: Page 27-33; 54pp; English.
 XX The protein product (R90528) of the FLT4 receptor tyrosine-kinase
 CC DNA sequence (t12068) is a novel marker for lymphatic vessels and
 CC some high endothelial venules. FLT4 is used to raise antibodies,
 CC esp. the monoclonal antibody produced by hybridoma DSM ACC 2210,
 CC useful in the diagnosis and therapy of e.g. inflammatory, infectious
 CC and immunological diseases, metastatic lymph nodes and lymphangiomas.
 XX Sequence 1298 AA:
 SQ

Query Match 95.0%; Score 6852; DB 17; Length 1298;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQRGAALCLRLWLCLGLLDGLVSGSYWPTPTLNITEESHVIDTGDSLSISCRGHPLEWA 60
 DB 1 mqrGaALCLrlwLclGLldglVsgSYwptPTlnITEESHVIDtgdsLSISCRghPLEwa 60
 QY 61 WPGAQAPATGDKDSEDVTGVVRCDEGTDRPYCKVLLLEHVHNDTGSYVCYKYIKARI 120
 DB 61 wpgAQAPAtgDKdSEdVTgVVrCDEgtDRpYcKVllLEhVhNDtGSyVcYkYiKaRi 120
 QY 121 EGTAAASYVFRDFEOPFTNKFDPTLLVNRKQAMWVCLVSIPLNVTLRQSSVLPDGG 180
 DB 121 egtAAAsYvfrDfEOPftNKfDPTllVnrKqAMwVclVsiPlNvTLrQssVlPDgg 180
 QY 181 QEVVMDRRGMLVSTPLLDHALVLCQETTWGDDFLSNPFLVHITCNELYDIOQLPRKSL 240
 DB 181 qevVMDrrGmLVstPlLDhAlVlcQETtwGDDfLSnPFVhITcNeLYdIOqlPRkSL 240
 QY 241 ELLVGEKLIVNCTVWAEFNSGVTFDMNDYPCQAERGKWPERRSQOHTTELSILTINHV 300
 DB 241 ellVgeKlIVnctVwAEfNSgvTFdMNdYpCqAeRGkWPerrSQOhtTElSIlTihNv 300
 QY 301 SOHDLGSYVCKANNGIQRFRETEVIVHENPFTISVEMWLGPILEATAGDELVKLPVLA 360
 DB 301 sohDlGSyVcKaNngIQrFrETeViVhEnPftISvEMwLgPiLeATaGdELvKlPvLA 360
 QY 361 YPPPEFOWYKDGKALSGRSHPHALVLKEVTEASTGYTTLALMNSAAGLRNLSLELVNV 420
 DB 361 yppPeFOWyKdGkAlSGrShPhAlVlKeVtEaSTgYtTLAlMnSaAGlRnLSlELvNv 420
 QY 421 PQQTHEEASSPSTYSRHSQALCTAYGVLPLSLQWHRPWPTCKMFAQNSLRQQQ 480
 DB 421 pqqTHEeAsspStYSrHSqALctAYgVlPlSLqWhrPWpTCKmFAqNSlRqqQ 480
 QY 481 DLMPQCRDRAVTTQDAVNP IESLDTWTEFVEGKNKTYSKLVTONANVSAMYKCVSNKV 540
 DB 481 dLmpqCRdRAvTTQdAVnPIeSLdTWtEFvEGKnKtYSkLVtONANvSAMyKcVsnKv 540
 QY 541 GQDERLIYFVVTIPQGFITESPSELLEGQVLLSCQADSYKYKYLRYRLNLTSLHD 600
 DB 541 gQderLIyFvVTiPqGfITeSpSeLlEGqVllSCqADsYkYkYlRyRLnLTslHD 600
 QY 601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLSIPRVAPCEHYVCEVODRKS 660
 DB 601 ahgnPlLLdCkNvHLfATpLaASlEEvAPgARhATLSlSiPrVAPcEHYvCEvODrKS 660
 QY 661 HDRKCHKYLSVQALAPRLTQNLDTOLLVNSDSLEMOCLVAGAHAPSIVWYKDERLLEE 720

Db 661 hdknchkYlsvqALeaprltqnltdllvnvsdslemqclvagahapsivwykderllee 720
 QY 721 KSGVDLADSNOKLSIORVREEDAGRYLCSVCNAKGCVNSSASVAVEGSEDKSGMEIVILV 780
 DB 721 ksgvDLadSnOKlSIORvREEdAGrYlCSvCNAkGCvNSSASvAVeGSEdKSGmeiVilV 780
 QY 781 GGVIAVFFVWVLLLIIFCNMRPAHADIKTGYLSIIIMDPGEVPLEDCEYLSYDASQWFF 840
 DB 781 ggViAvFFvWvLLlIFcNmRPAhADIKtGYlSIIIMdPGEvPLEdCEyLSyDASqWff 840
 QY 841 PRERHLGRVLGYGAFGKVVVEASAFGIHKGSSCDTVAVVRLKEGATASEHRLMSELKITL 900
 DB 841 prerHlGrVlGyGafGkvVeASafGIhKgSScdTVaVvRLKEgATaSEhRlMseLkiL 900
 QY 901 IHIGNHLNVNLLGACTKPGQPLMWIVFECYKGNLSNFRAKRDAPSPCAEKSPQGRGF 960
 DB 901 ihIghNhlNvNlLGacTKpGqPlMwIVfEcYkGNlSNfRAkRDAPSPcAEKSPQGRgf 960
 QY 961 RAWVELARLDRRRPGSSDRVLFAFESKTEGGARRASPDQAEEDLWLSPLTMEDLYCSFQ 1020
 DB 961 rawVeLaRlDRrRPgSSdRvLFAfESkTEgGARrASPDqAEEDlWlSPlTMEdLyCSfQ 1020
 QY 1021 VARGMEFLASRCKIHRDLAARNILLSESVDVVKTCDFGLARDIYKDPDYVRKGSARLPLW 1080
 DB 1021 varGMEfLaSRcKIhRDLAARNillSEsVDvVKtCdFGLaRDIyKdPDyVRkGSARlPlW 1080
 QY 1081 MAPESIFDKVYTTQSDVWSFGVLLMEIFSLGASPPYGVQINEEFQORLDGTRMRAPELA 1140
 DB 1081 mapESiFdKvYttQsdVwSfGvLLmEiFslGASpPYgVqINeeFqORldGTRmRApELa 1140
 QY 1141 TPAIRRIMLNCHSGDKARPAPSELVEILGLDLOGRGLOEEVEVCMAPRSSOSSEGSFS 1200
 DB 1141 tPaIRrImLnChSGdKArPaPseLVEilGLdLOGRGLOEEvEVCmAPrSSosSeGSfs 1200
 QY 1201 QYSTMALHIAQDAEDSPSLQRHSLAARYNWNVSPFGCLARGAETRGSSRRMKTTEEPFM 1260
 DB 1201 qyStMaLhIaQDAEdSPslQRhSLaARyNwNVsPFgCLaRGaETRGssRRmKtTEEPfM 1260
 QY 1261 TPTTVKGSVDNOTDSCHWLASEEPEQIESRHRQESGF 1297
 DB 1261 tPtTVkGSvDnOTdSCHwLASEePEqIEsRHrQESgf 1297

RESULT 3
 W80597
 ID W80597 standard; Protein; 1355 AA.
 XX
 AC W80997;
 XX
 DT 12-APR-1999 (first entry)
 XX
 DE Human receptor tyrosine kinase KDR.
 KW KDR: receptor tyrosine kinase; human; signal transduction;
 KW mitogen; neoangiogenesis; angiogenesis; diabetic retinopathy;
 KW breast cancer; brain cancer; inflammation; rheumatoid arthritis;
 KW psoriasis; contact dermatitis; hypersensitivity; antagonist;
 KW therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 780..1386
 FT /note= "intracellular domain, from about amino acid
 FT 780-795 to about amino acid 1175-1386"
 XX
 PN W09858053-A1.
 XX
 PD 23-DEC-1998.
 XX
 PF 17-JUN-1998; 98WO-US12569.
 XX

18-JUN-1997; 97US-0050962.
(MERI) MERCK & CO INC.
Kendall RL, Mao X, Tebben A, Thomas KA;
WPI; 1999-095333/08.
N-PSDB; V99829.
Human receptor tyrosine kinase protein, KDR - useful e.g. to screen
for antagonists useful to treat diseases involving neovascularization
e.g. diabetic retinal vascularization, cancers
Claim 2; Fig 2; 69pp; English.
This is the amino acid sequence of a novel receptor tyrosine kinase,
termed KDR, that is expressed on human endothelial cells. KDR is
activated by vascular endothelial growth factor and mediates a
mitogenic signal. It is implicated in clinical neovascularization.
The amino acid sequence was deduced from an isolated KDR cDNA (see
V99829). The predicted sequence has differences from the previously
published KDR sequence at positions 498 (Ala to Glu), 772 (Thr to
Ala), 787 (Gly to Arg), 835 (Asn to Lys), 848 (Glu to Val) and 1347
(Thr to Ser), producing a protein predicted by computer modeling to
have higher activity and functionality. The invention also relates
to recombinant vectors and recombinant hosts which contain a DNA
fragment encoding human KDR, a DNA fragment encoding the
intracellular portion of KDR with or without a membrane anchor
sequence, purified forms of associated human KDR, and human mutant
forms of KDR. KDR, fusion proteins or fragments can be used in
assays to identify antagonists and agonists of human KDR (claimed).
Antagonists of KDR useful for treating diseases involving
neovascularization e.g. diabetic retinal vascularization, cancers
(e.g. brain, breast, etc.) and forms of inflammation e.g. rheumatoid
arthritis, psoriasis, contact dermatitis and hypersensitivity
reactions. The protein is also useful for produce antibodies that
can be used to measure human KDR levels and/or activity.

Query Match 38.2%; Score 2753; DB 20; Length 1355;
Best Local Similarity 44.1%; Pred. No. 9.2e-193;
Matches 605; Conservative 204; Mismatches 476; Indels 88; Gaps 26;
7 LCLRLWLCGLDGLVS--GYSTPTPLNTIESHVIDTGDLSISCRGHPLEWAPGA 64
6 lavalwlcetraasvslpvsldlprlsiqdkdltikantlqitcrgardldwlpnn 65
65 QEAPATGDKDSEDGVVVRDC-EGTDARPYCKVLLHEVHANDTGSVYCYKIKARIEGT 123
66 qsg-----seqrvevrecsdgl----fckltipkvigndgkaykcfy-----etd 108
124 TAASVYVVRDFEOPFINKPT-----LLVNRKDAMVCLVSLIPGLNLT--RSOSSV 175
109 lasviyvyvqdyrsfpiaasvqhgvyvitenkntkvipclgslnvslcarypekr 168
176 LMPDGOEVVWDRRGMVSTPLLDALYLLOCETTWGQDQLSNPFLVHITGNELYDIQLL 235
169 fvpdgnrswskkgftipsymisagmveaceakindesysqymyvvvgyrydyvls 228
236 PKSLLELVGEKLVNCTVMAEFNSGVTFDWDYPCGAERCKWYPERRSQOHTHLS--- 292
229 pshgielsvgeklvinctartelnvgidfnweypssqhkhklnrdlklqsgsemkkl 286
293 STLTTHNYSOHDLSGVYCKANNGIORPRESTEVIHNPFLISVEWLKGPILAEATAGDELV 352
289 stltldgvrtsdgglytcaasgintkknstfvrvehkpfvafgsgmesiveatvg-erv 347
353 KLPVKLAAYVPPPEFQWYKDGKALSGRH---SPHALVKEVTEASTGTYTTLALWNSAAGLR 409
348 ripakylgypppeikwngknglplesnhtikaghlvtimevserdtnyvtiitnpskek 407

QY 410 KNISEILVNVPPQIHEKEASSP-SIYSRHSROALICTAYGVPLPLSIOHHRPWTPTCKM 468
DB 408 qshvsvsvvvpqigeekslispsvdygttctctvvaippphhhiwqgleec-- 465
QY 469 FAURSLLRRQOQDLMPOCRDMRAVTTODAVNPITESLDTTEFVEGKNKTVSKLVIQANV 528
DB 466 -anepsqavsvtnpyp-ceedrsvedfggnklevknknqfalliegnkktvstivgaanv 523
QY 529 SAMYKCVSNKVGQDERLIYFYVTTIPDGFITIESEKPESELEGOPVLLSQDASYYEHL 588
DB 524 salykeavknvrgervisfhvtrgpe---itlqdmqpteqesvslwtadstfnel 580
QY 589 RMYRLNLSTLHDAGNPLLLDCKNV----HLFATPLAASLEEVAPGARHATLSLIPRVA 644
DB 581 twykglpqpipihvgelptpvcnkldtlwknatmsfnsndi-----lmeiknas 632
QY 645 PEHEGHVCEVQDERRSHDKHKYKLVSVQALEAPRLTONLTDLLVNVSDSLEMOCLVAGA 704
DB 633 lqagdyvclaqdrktrhcvvqlcvlervaptlgnlenqtsiesiesvctasgn 692
QY 705 HAPSIWYKDERLLEKSGVDLADSNOKLSIORVREDAGRYLCSVCNAKCVNSSASVA 764
DB 693 pppqlmwfkdneltvedsgivlkdgnrnltrrvrkedeglytcqacsvlgcaveaffi 752
QY 765 VEGSEDKGSMIEIVILVGTAVFFWVLLIFCNMRRPARHADIKTGYLSLIMDPGEVPL 824
DB 753 legaqeknlleilvgtavlamffllvllvtkranggelktgylsvmvpdpdelpl 812
QY 825 EEOCEYLSYDASQWEPFRERLHRLVGLYGFAGVKNVASAFGIHKGSCDTAVAKLKEG 884
DB 813 dehcrlpydaskwefprdklgrgafgvideadafgdktdtctrtvavmkikeg 872
QY 885 ATASEHRALMSELKTLIHGHNHLNVNLLGACTKPOGLVWIVFVCKYGNLSNFRKRD 944
DB 873 ahshhralmelkllhghnlvnlvgactkpgplmvivefckfnglnstylrskrn 932
QY 945 AFSPCAEKSPEQGRFRA---MWELARLDRRRPGSSDRVLFAFESKTEGGARRS----P 997
DB 933 efvpvktkg---arfrqgkdyvgaipvdlkr--rldstssqssasgfvveeksldve 986
QY 998 DOEAE-EDLWLSPLTMEDLVCSYFQVARGMEFLASRKC IHRDLAARNILLESVDVYKICDF 1056
DB 987 eeeapepdykdfiteltchlycsfvaqkmefflasktrchrdlaarnillseknvktcdf 1046
QY 1057 GLARDIYKDPDYVRKGSARLPLKHMAPPESIFDKVYTTOSDVNSFGVLLWEIPLSGASYP 1116
DB 1047 glardiykdpdyvrkgdarlplkwmappetifdrvylgdsvsgfllwelfsigasyp 1106
QY 1117 GVOINEEFORLDRDCTWRAPELATPARRIMLNCWSDGPKARPAFSELVETLGLLOQR 1176
DB 1107 gvkldeefcrrlkegrmrpadyttcmvqtmldcwhgsepqrptfselfvehlgnllgan 1166
QY 1177 GLQEEVEVCWAPRS--SQSSEEGSFQVSTMALHIAQADSPSPQLRHSLAARYNWS 1235
DB 1167 aqdgdykdyivlpisetlmeedsgislptspvscmeveevcdp-----kfydntag 1218
QY 1236 FPCCLAGARETRGSSPMKTFEEFPM-TPPTYKGSVDNOTDSGMVLASEEFOIESRHQOE 1294
DB 1219 isqylqnskrksrpsvsktfedipeevkvpdpndqdsqgmviaseekkletrtklis 1278
QY 1295 SGFSCGKQPQNVAVTRAHPDSQGRRRRPPRGARGQGVFNSEYSELSEFSEED 1347
DB 1279 psfgmvpak---sresvasegsnq-----tsqyqsgyhsddtdttvyssee 1322
RESULT 4
K26999
ID R26999 standard; Protein: 1356 AA.
XX
AC R26999;
XX
DT 20-MAY-1998 (first entry)
XX

Novel type III RTK encode by the KDR gene.

Receptor tyrosine kinase; vascular endothelial cell growth factors; cancer; tumour; diagnosing; monitoring.

Homo sapiens.

Location/Qualifiers

46 [potential N linked glycosylation site.](#)

99

96

143

158 potential N-linked glycosylation site

/note= "potential N-linked glycosylation site
245

/note= "potential N-linked glycosylation site"

/note= "potential N-linked glycosylation site"

574
/note= "potential N-lined glycosylation site"

395 /note= "potential N-linked glycosylation site"

511

523

580

613

/note= "potential N-linked glycosylation site
631

/note= "potential N-linked glycosylation site"

/note= "potential N-linked glycosylation site"

potential N-linked glycosylation site"

/64.:788
/note= "putative membrane spanning region"

1..763

contains 21 cysteine residues"

```
/note= "ATP binding site"
```

WQ9214748-A

03 - SEP - 1963

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

44-FEB-1991; 9103-0037230.

(AMCY) AMERICAN CYANAMIDE

Carrion ME, Terman B

WPI; 1992-316115

N-F3DB, Q202/2.

DNA encoding type III receptor diagnosing the onset of cancer

claim 3. Fig 7. 101mm. English

This sequence represents a novel type III receptor tyrosine kinase encoded by the KDR gene. It was decoded from the appropriate cDNA, and its predicted amino acid sequence contains several features which demonstrate that the novel gene is a type III RTK. A hydropathy plot of the predicted amino acid sequence indicates a single membrane spanning

Db	748	veaffliiegaqektuleiilivgtvtviamffwllilviltvkrangeltgylslvmd	807
Qy	819	PGEVPLEEOCEVLSYDASQWEPFRERLHGRVLGYGAFKGVVEASAFGIHKSGSCDTAVV	878
Db	808	pdeidhecerlpydaskwefdrlnlglprgafgqeieadagidktatctvav	867
Qy	879	KMLKEGATASEHRAIMSELKILIHIGHNLNVNLLGACTKPGQLVIVVEFCYKGNLSNF	938
Db	868	kmikegathsehraImseIkiliihiighlnvnlilgactkpggplmvivefcfgnlsty	927
Qy	939	LRAKRDAFSPCAEKSPQORGFRA---MVLEARLDORRRPGSSDRVLFAFSKTEGGARRA	995
Db	928	lrskrmefvpytkg----arrrgqkdyvgaipvdllkr--rldsitssqssassgfveek	981
Qy	996	S-----PDQEA-EDLMLSLPTMBDLVCYSFOVARGMEFLASRKCITHRDIAARNILLSES	1050
Db	982	slsdsveeeapeadlykdfitlthlicysfvakgmeflasrkcIhrdlaarnillseknv	1041
Qy	1051	VKICDGLGLARDLYKQPDYVRKGSARLPLKWNAPESIPDKVYTTQSDVMSFGVILLWEIFSL	1110
Db	1042	vkicdglglardlykdpdyvrkgdaripikwmapetifdrvytiqsdvmsfgvillweifsl	1101
Qy	1111	GASPYPGVQINEEFCQRLRDGTRMRAPELATPATRRIMLNCWSGDPKARPAFSELVEILG	1170
Db	1102	gaspypgvkideefcrikegtrmradyttipemyqtmldcwhgpepsqrptfselvehlg	1161
Qy	1171	DLLOGRGLQEEVEVCMAAPS-SQSSEEGSFQSVSTMALIAQADAEDSPSLORHSLAAR	1229
Db	1162	nllqanaqdgkdyivlipsetlmeedsGISLTPVSCMEEEVCDP-----kfh	1213
Qy	1230	YNNWVFPQCLARGAETRGSRMKTFEEFPM-TPTTYKGSVDNQTDSGWLASEEPEQIE	1288
Db	1214	YDNTAGISGLNSKRSKRPVSKTFEDIPLEEPVKVIPDDNQTDSGWLASEEIKTLE	1273
Qy	1289	SRHRQSGFSGKPGQCNVAVTRAHPDSQRRRRPERGARGQVFPYNSEYGEIASEPSEED	1347
Db	1274	drtklspsfggmvpsk---sresvasegsnq----tsqyqsyghsdtdttivyssee	1323
RESULT	5		
W59275			
ID	W59275	standard; Protein: 1356 AA.	
XX	AC		
XX	W59275;		
XX	DT	27-AUG-1998 (first entry)	
XX	DE	Human KDR protein.	
XX	DE		
KW	KW	Kinase insert domain containing receptor; KDR; screening; inhibitor;	
KW	KW	vascular endothelial growth factor; VEGF; angiogenesis; treatment;	
OS	OS	cancer.	
XX	XX	Homo sapiens.	
XX	XX		
FH	Key	Location/Qualifiers	
FT	Modified-site	46	
FT	FT	/label= N-glycosylated	
FT	FT	/note= "putative"	
FT	Modified-site	66	
FT	FT	/label= N-glycosylated	
FT	FT	/note= "putative"	
FT	Modified-site	96	
FT	FT	/label= N-glycosylated	
FT	FT	/note= "putative"	
FT	Modified-site	143	
FT	FT	/label= N-glycosylated	
FT	FT	/note= "putative"	
FT	Modified-site	158	
FT	FT	/label= N-glycosylated	
FT	FT	/note= "putative"	
FT	Modified-site	245	
FT	FT	/label= N-glycosylated	

FT		/note= "putative"
Modified-site	318	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	374	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	395	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	511	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	580	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	613	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	631	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	675	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	704	
FT		/label= N-glycosylated
FT		/note= "putative"
Modified-site	721	
FT		/label= N-glycosylated
FT		/note= "putative"
Region	765..789	
FT		/note= "putative membrane spanning region"
XX		
US5766860-A.		
PN		
XX		
PD	16-JUN-1998.	
XX		
PF	25-FEB-1997;	97US-0810116.
XX		
PR	23-NOV-1992;	92US-0930548.
PR	25-FEB-1997;	97US-0810116.
XX		
PA	(AMCY) AMERICAN CYANAMID CO.	
XX		
PI	Carrión ME.	Terman BI;
XX		
DR	WPI; 1998-361682/31.	
DR	N-PSDB; V34763.	
XX		
PT	Screening assay for vascular endothelial cell growth factor antagonists - using recombinant cells expressing receptor protein	
XX		
PS	Claim 1; Fig 7A-M; 51pp; English.	
XX		
CC	This sequence represents a novel human growth factor receptor, kinase insert domain containing receptor or KDR. This receptor is capable of binding to the vascular endothelial cell growth factor, VEGF and is used in a screening assay which identifies compounds that inhibit VEGF action on KDR. Such compounds which inhibit binding of VEGF to the KDR may inhibit angiogenesis and thus be useful for treating cancer.	
CC		
Sequence	1356 AA;	
Query Match	38.2%;	Score 2753; DB 19; Length 1356;
Best Local Similarity	43.9%;	Pred. No. 9.3e-193;
Matches	60%; Conservative 205; Mismatches 480; Indels 88; Gaps	
OY	1 MORGALRLRLWLCLGLDCLIVS--GYSWTPPTLNITEESHVIDTGSLSSCRQHPLE 58 : : 1 1 : : : : : : : : : : : : : 1 meskvllavallvcetraasvglpvsaldiprlisigkdlitkanktllgltcrggrddl 60	
Db		

QY 59 WAPGAQAEAPATGDKDSEDGTGVRDC - EGTDAAPYCKVLLHVEHNDTSGVVCYKIK 117
Db 61 WLPANQSG -----seqrvetcesdgl-----fckltipkvlgndtgaykcyf--- 106
QY 118 ARIEGTTAASSYFVDFEQEP INKPD T-----LLVRKDAHWVPCLVISPLNWL-- 169
Db 107 ----etdlasviyydyrsfpiaasvqhgvyvitenkntvtpcpgslnlnvsica 163
QY 170 RSOSSVLPDQGEVWDDRRGLMSTPLLDALYLQCTETGQDQDFLNPFLVHITGNE 229
Db 164 rypekrfpvdpnriswskdktfipsymisyagmvfcaekindexysqymyvvvgyri 223
QY 230 YDIQLPRKSLLELGEKLVNCTVMAFENSGVTDFDWDYPGQAEGRKWVPERRSQOHT 289
Db 224 ydvvisphglsvegelvlnctartelnvgidfnweypsskhhqkklvnrldktqs 283
QY 290 ELS---SILTIHNSQHDLSYVCKANNGIORFRESTEYIVHENPFISVWELKGPILEAT 346
Db 284 emkfkfistldgvtresdgltytcaassglmckknstfvrhvekpfaavgsgmeslveat 343
QY 347 AGDELVKLPVLAAYPPPEFQWYKDGKALSGRH---SPHALVLKVEYTEASTGTTLALWN 403
Db 344 vg-ervripakylgyppepkelwkniplesnhtikaghnvltmevserdtnyvtvllcn 402
QY 404 SAAGLRNLTSLVNVNPPQIHEKEASSP-SIYSRHSRQALCTAYGVPLPLSIQWHWRP 462
Db 403 pisekqshvslvvyppqigekslisvdsyqygttqltctvyaipphphihwywl 462
QY 463 WTPCKMFAORSRRRQOQDLPCQCRDWRVAVTQDANPIESLDTWTFEVEGKNKTVSKVL 522
Db 463 eeeec---aneqsdavsvtnpy-pceewrsvedfggnkievknqfallgknktvstiv 518
QY 523 IQANVANSYKCVVKNKGDERLIYFYVTTIPDGFTIESKPEEELLEGOPVLLSCQADS 582
Db 519 lqanvnsalykcaavknvgrgvisfhvtrgpe---itlqdmqptequesvslwctadr 575
QY 583 KYEHLRWYRLNSTLHDANGNPLLLDCKNV-----HLFATPLAASLEEVAPGARHATLSL 638
Db 576 stfenltwylgqpipihvgelptpvcnkldtlwklnatmsfnstndi-----lim 637
QY 639 SIFRAPEHGHVCEVQDERSHDKHKYLSVQALEAPRLTONLTDLVNVSDSLEMQ 698
Db 628 elknasldgdgyvclaqdktrkrcvvrqltclervaptitngntqtsigesievs 687
QY 699 CLVAGAHAPSIVNWKDERLEERSGVDLADSNOKLSTQVRREEDAGRYLCSVCNARKGCYN 758
Db 688 ctasgnppqimfwkdnnetivedgvlvkgdnrltirrvkedeglytcqacsvlgcak 747
QY 759 SSASVAVESDEKSGMEIVTLVGTGVITAVFFWVLLLLIFCNMRRPAHADIKTYLSIIMD 818
Db 748 veaffilegaqektneililvgttviamffvllvllgtvkranggeliktgylsivmd 807
QY 819 PGEVPLEEQEYLSYDASOWEPRERHLGLRGVYGAGFVKEASAFGIHKGSSCDTVAV 878
Db 808 pdeipldeheerlpydaskwefprdrinlgkplgrgafgeieadafgtdktatcrvav 867
QY 879 KMLKEGATASEHRALMSLEKILHIGHNLVNVLGACTKPOGLMIVIVEFCYKGNLSNF 938
Db 868 kmikegathsehralmseklilhighlnvnlvgactkpggplmvivefcykgfnlsty 927
QY 939 LRAKRAFSPCAEKSPQRCRFRPA---MVELARLDRRRPGSSDRVLFARSKTEGGARRA 995
Db 928 lrskrnefvpyktkg-----arfrgqdyvgaipvdlkr---rldsitssqssasgfwcek 981
QY 996 S-----PDQEA-EDLWLSPLTMEDLVCSYFQVARGMEFLASRKCIHRDLAARNILLSSESV 1050
Db 982 slsdveeeceapedlykfltlehlicysfvcqkmgfslasrkcihrdlaarnillseknv 1041
QY 1051 VKICDFGLARDIYKDPYVRKGSARLPLKWMAPESIFDKYVTTQSDVWSFGVLWEIFSL 1110
Db 1042 vkicdfglardiypkdpvkrkgdarlplkwmapetifdrvytlqsdvwsfgvllweifsl 1101
QY 1111 GASPPYGVQINEECQRLRDRGTRMRAPELATPAIRRIWMLNCWSDCPKARPAFSELVEILG 1170

Db 1102 gaspypgvkiideeferrlkegrtmrapdyttmpemqtmldcwhgepsqrptfseivehlg 1161
QY 1171 DLLQGRGLQEEEEVCMAPRS-SQSSEEGSFQSVTMAHLHIAQADAEDSPPSLQRHSLAAR 1229
Db 1162 nllqanaqgdgkdyivlpiselimeedsgisipcsyvsmeeeevcdp-----kfh 1213
QY 1230 YYNWVSFPCCLARGAETRGSSRMKTFEEFPM-TPPTYKGSVDNQTDQSGHVLASEEFAQIE 1288
Db 1214 ydnlagisqylonskrksrpsvktfedipleepevkvipddoqldsgmvlaseelktle 1273
QY 1289 SRHROESGFCCKGPCQNVAVTRAHPDSCRRRRRPPRGAGGOVFNSEYSELSEPSBED 1347
Db 1274 drtklspfsfgmvpksk-----sresvasegsnq-----tsyqsgyhsddtdttvyssee 1323
RESULT 6
R44996
ID R44996 standard; Protein: 1367 AA.
XX
AC R44996;
DT 27-JUN-1994 (first entry)
XX Murine flk-1 receptor protein tyrosine kinase.
DE Receptor protein tyrosine kinase; PTK family; foetal liver kinase;
KW hflk; primitive; totipotent; haematopoietic cell; stem cell;
KW proliferation; stromal cell.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= signal sequence
FT /note= "hydrophobic leader"
FT Protein 20..1367
FT /label= flk-1
FT /note= "mature protein"
FT Domain 20..762
FT /label= extracellular_domain
FT Region 763..784
FT /label= transmembrane_region
FT Domain 785..1367
FT /label= intracellular_catalytic_domain
XX
PN US270458-A.
XX
PD 14-DEC-1993.
XX
PF 02-APR-1991; 91US-0679666.
XX
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0906397.
PR 12-NOV-1992; 92US-0975049.
PR 19-NOV-1992; 92US-0977451.
XX (UYPR-) UNIV PRINCETON.
XX
XX Lemischka IR:
XX
DR WPI; 1993-405021/50.
DR N-PSDB; Q53504.
XX
XX Isolated nucleic acid molecules of hematopoietic stem cell
XX receptor flk-2 - encoding mammalian receptor protein tyrosine
XX kinases expressed in primitive haematopoietic cells
PS Disclosure: Fig 2; 60pp; English.
XX

CC Nucleic acid sequences coding for murine flk-2 and specified
CC subfragments of it are claimed. The murine flk-1 coding sequence
CC (i.e. Q3504) is also disclosed. The flk polypeptides are
CC receptor protein tyrosine kinases which are expressed only in
XX primitive haematopoietic cells.
SQ Sequence 1367 AA;

Query Match 37.0%; Score 2665; DB 14; Length 1367;
Best Local Similarity 42.9%; Pred. No. 2.6e-186;
Matches 587; Conservative 205; Mismatches 486; Indels 90; Gaps 24;

QY 1 MQRGAALCLRLCLGLDGLV--GYSMTPTPLNTEFESHVITDGLSLISCRGHPLE 58
DB 1 mesgllavalfcvetraasvypdflhpklsktdiltillatltqitcgrqdlid 60
QY 59 WAMPQAEAPATGDKSDGTVVDRDCEGTADPYCKVLLHEVHANDTSYVCYKYKA 118
DB 61 wlpnaq-----rdseerlvltcgggds-ifckltltiprvvngndgaykcsrdv-- 110
QY 119 RIETTAASSYVFVRDFEOPFINKPDT-----LLVNRKDMWVPCLVSTIPGLNVTL--R 170
DB 111 ----diastvyvdyrsfiasvsdqhgivytentkntkvipcrgsinlnvsicar 166
QY 171 SOSVLPDQGVVDDRRGMLVSTPLLDALYLQGETTWGDQDFLSNPFLVHITNELY 230
DB 167 ypekrfvpdgnrswdselgtlpsymisyagmvfcaekindetyqslmyivvvgyriy 226
QY 231 DIQLPRKSLLELLVCKLNCVWAEFFNSGVTFDDNDYPCQKQERKWWPERSQOQTHTE 290
DB 227 dvlisppheleisagekivincartelnvgldftwhspkskshhkvlnrdvkpfpgvt 286
QY 291 LS----SILTIHNSQHDLSYVCANNGIQRRFESTEVIVHENPFISVEMLKGPILEATA 347
DB 287 akmlstltiesvtksdqgeytcvassgrmknrtfrvrvhtkpfiafgsgmksiveatv 346
QY 348 GDELVKLPVLAAYPPPEFOWYKDGKALSORHS---PHALVLKEVTEASTGTITLALWNS 404
DB 347 gsg-vripvkylsyppadikwyngnrpliesnytmlyvgdelitmevterdagnvltitnp 405
QY 405 AAGLRNLSLELVNVPQIHEKEASPS-SIYSRHSRQALTCYAGVPLSLSTQWHRPW 463
DB 406 ismekqshmslvvvpqigekallspmdsyqygmqltictctyanpplhhqiywqle 465
QY 464 TPKMFAQRLRRRQOQDLMPQCRDRAVTTQDANPIESLDTWTFEVECKNTVSKLVI 523
DB 466 eacsy-----rpgqtsyackewrhvedfqqgnkievknqyallegknktvstlvi 517
QY 524 QNAVSAWYKCVSNVKGQDERLIYFYVTTIPDGFTIESKPSLELGGOPVLISCOADSY 583
DB 518 qanvsalykea inkagrvvisfnvirgpe---ltvqpaagqteqsvislctadrn 574
QY 584 KYEHLRWYRLNLSLTHDAHGNPLLLCKNY----HLFATPLAASLEEVAPCARHATLSLS 639
DB 575 tfenltkykgsqatsvhmgesltvpcnkldalwklingtmfsnstndi-----liva 626
QY 640 IPRVAPHEHYCEVDORSHDKHKYLSVQALEAPRLTONLTDLLVNVSDSLEMOC 699
DB 627 fgnaslqddqdyvcsaqqdktkrhvklvqlillemapmitnlenqtttigittevtc 686
QY 700 LVAGAHAPSVWYKDERLLEKESVDLADSNOKLSLQTVREEDAGRYLCSCVNAKGCVNS 759
DB 687 pasgnptphltwfknetlvedsgvlrdgnrnltrrvrkedgglytcaqcnvlgcara 746
QY 760 SASVAVGSEDKSMEIVILVGTCTAVTFWVLLLLIFCNHRRPARADIKTYGLSIIMDP 819
DB 747 etlfiiegagektulevliivgtaviamffvllvltvrvtkranegeltgylstfwdp 806
QY 820 GEVPLEQCEYLSYDASOWEPPRERHLGVRVLYGAFGKVVESAFGIIHGKSSCDTVAVK 879
DB 807 delpldrcerlpydaskewfprdrkrlgkplrgtgrgavgieadafgdktdactkvavk 866

QY 880 MLKEGATASEHRALMSELKILIHIGNHLNVNLLGACTKPOGLMVIIVEFCKYGNLSNPL 939
DB 867 mkegathsehralmselekilighhlnvnlngactkpgplmvivefskfgnlstyl 926
QY 940 RAKRDAFSPCAEKSPQGRFPA----WVELARLDRRPGSSDRVLFARFSKTEGGARRA 995
DB 927 tgrknefvpykskg----arfrgkdyvgeis-vdlkr--rldsttsqssassgfveek 979
QY 996 S-----PQEAEDLWLSPLTMEDLVYCSFOVARGHEFLASRKCIHRDLAARNILLSSEV 1050
DB 980 slsdveeeaseelkydfllehllicysfvakgmeflastrkchrdlaarnillseknv 1039
QY 1051 VKICDFGLARDIYKDPDYVRKGSARLPLKWMAPESIFDKVYTTQSDVMSFGVLLWEIFSL 1110
DB 1040 vxicdfglardiypdyvrkgdarlpikwmapetifdrvytiqsdvsvfllweifsl 1099
QY 1111 GASPPYGVQIINEFCORLDRGTMRAPELATPAIRIMLNCWSGDPKARPAFSELVEILG 1170
DB 1100 gaspygvkideefrrlkegtrmapdyttptemygtmldcwhepndqprpsfseivehlg 1159
QY 1171 DLLQGRGLQEEVEECVMAPRS--SQSSEEGSFQSVSTMALHIAQDAEDSPSLQRHSLAAR 1229
DB 1160 nllqanaqqdgkdyivlpmssetlsmeedsglsiptspvscmeeeevcdp-----kch 1211
QY 1230 YNNVVSFPCLARGAETRCSSRMKTFEFPFM-TPPTYKCSVDNQTDSCHVLASEEPEQIE 1288
DB 1212 ydntagishylgnskrkprsvskttfediplateevkviipdsqtdsgmvlaseelktie 1271
QY 1289 SRIHQESGFSCKPGPONAVTRAHPDSQGRRRRPERGARGQGVFNSE 1336
DB 1272 drnkispsfg----gmppsksrsvasegsnq-----tsyqsgyghsd 1310

RESULT 7
R31377 ID R31377 standard; Protein: 1367 AA.
XX AC R31377;
XX DT 25-JUN-1993 (first entry)
XX DE Human flk-1.
XX KW Human; receptor; protein; tyrosine kinase; pTK; primitive; mammalian;
KW hematopoietic cell; ptk; mature; mhc; fetal; liver kinase 2; flk-1;
KW liver; spleen; thymus; adult; brain; marrow; thymocyte; subset;
KW multipotential; T-lymphoid; lineage; stomach; kidney; lung; heart;
KW intestine; muscle; lymph node.
XX OS Homo sapiens.
XX PN WO9300349-A.
XX PD 07-JAN-1993.
XX PF 26-JUN-1992; 92WO-US05401.
XX PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 02-APR-1992; 92WO-US02750.
XX (UYPR-) UNIV PRINCETON.
XX PA Lemischka IR;
XX PI WPI; 1993-036323/04.
XX DR N-PSDB; Q35251.
XX Nucleic acid encoding receptor protein tyrosine kinase - allows
PT development of ligands to stimulate proliferation and/or
PT differentiation of mammalian haematopoietic stem cells
XX

PS Claim 10; Fig 2; 78pp; English.

CC This sequence represents a human receptor protein tyrosine kinase which
 CC belongs to a new functional class of protein tyrosine kinases (ptks).
 CC However, this ptk is not in the same class as fik-2. ptk in the same
 CC class as fik-2 are expressed in primitive mammalian hematopoietic
 CC (pHC) cells but not in mature hematopoietic cells (mHC). This gene
 CC is expressed in more mature hematopoietic cells. This protein is an
 CC example of a receptor ptk and is called fetal liver kinase 1 (fik-1).
 CC fik-1 is expressed in fetal liver, spleen, thymus, brain, stomach,
 CC kidney, lung, heart and intestine, and adult brain, kidney, heart,
 CC spleen, lung, muscle, marrow and lymph nodes.

XX Sequence 1367 AA;

Query Match	37.0%; Score 2665; DB 14; Length 1367;
Best Local Similarity	42.9%; Pred. No. 2.6e-186;
Matches	587; Conservative 205; Mismatches 486; Indels 90; Gaps 24;
QY 1	MORGAALCLRLWLCGLLDGLVS--GYSWTPPLNTEESHVDTGDSLSISCRGQHPLE 58
DB 1	mesgllavalwfcvetraasvlgpdlhpkxistqkdltilantctlgtrggrld 60
QY 59	WAWPGADAPATGDKDSEDGTGVVRDCEGTDPAPYCKVLLLEHVNADTGSVCYCYKAIKA 118
DB 61	wlpnaq-----rdseervlvecggds-ifeaktitprvvgndtgayksyrdv-- 110
QY 119	RIEGTTAASSVYFVRDFEQPFINKPDT-----LLVNRKDMWVCLVSLIPGLNVL--R 170
DB 111	-----diastvvyvdyrpfiasvdsqhgivitenkntkvipcrgislnvslcar 166
QY 171	SQSVLWPDQGVVDDRRGMLVSTPLLDHALYLCQETWGDQDFLSNPLVHITGNELY 230
DB 167	yepkrfpdgnrtdseigtflpsymisyagmvfcekndetygsimyyvvvgyrly 226
QY 231	DIQLPRKSLLELVKLVNCTVWAEFNSGVTFDDYDYGKQAEKGVKWPERRSOOTHT 290
DB 227	dvilspheielesageklvinctartelnvgldftwhspsskshhkvnrdrkpfpgtv 286
QY 291	LS-----SILTHNVSOHDLGSLVYCKANNQIQRFRETEVIVHENPFISVEMKGPIL 347
DB 287	akmflstltvtskdsqgeytcvassgmikrntfrvrvtkpfafgsgmkslveatv 346
QY 348	GDELVKLPVLAAYPPPEFQWYKDGKALSORHS---PHALVLKEVTEASTGYTLALWNS 404
DB 347	gsq-vripkylsyppadikwyrngripiesnytmvgdeltimevterdagnyvtltnp 405
QY 405	AAGLRNLSLELVVVPQTHEKEASSP-SIYSRHSRQALCTAYGVPLPLSTQWHRPW 463
DB 406	ismekqshmvsvvvpqigekalispmdsyqygtemqtltctvyanpphlhiqwywle 465
QY 464	TPCKMFAQSLRRQQDLMPQCRDRAVTTQDAVNPISLDTWTTEVECKNTVSKLVI 523
DB 466	eacy-----rpdqtsyackewrhvedfggnkievknqyaliegknktvtlvi 517
QY 524	QNAVSAHYKCVSVNKGODERIYYVVTTPDGFTIESKPSBELLEGQPVLLSCQADSY 583
DB 518	qaanvsalykceainkagrgervishvirgpe---itvqpaapteqsvsllctadrn 574
QY 584	KYEHLYRWLNLSLTHDAHGNPLLLDCKNV---HLFATPLAASLEEVAPGARIATLSLS 639
DB 575	tfnitwyklgsqatsvhnagesitpvcnkldalwkingtmfsnstndi-----lva 626
QY 640	IPRVAPHEHYVCEVQDRSRHDKCHKKYLSVQALEAPRLTQNLTLVNVVSDSLEMO 699
DB 627	fqnasldqdgdyvcsaqdktkrhclvklqllilermapimignlqntttigetievic 686
QY 700	LVACAHAPSIVWYKDERLLEESGVDLADSNOKLSIQRVREDAQRYLCSVCAKOCVNS 759
DB 687	pasgnptphitwfkdnctivedsgivldrgnrlttrvrkedgglytqcacnvlgcara 746
QY 760	SASVAVGSESDKGSMEIVLVGVGTAVFFWVLLLIIFCNMRRPAHADIKTGYLSIIMDP 819

Db 747	etlfllegaqektnleviilvgtavlamfvlvllvtrvkranegeklktgylsvmdp 806
QY 820	GEVPLEEQCEIYSYDASOWEFPRERLHIGVLCYGAFAKVVVEASAFGIHKSGSDTVAVK 879
DB 807	delpidercerlpydaskwefpdrklgkplgrgafgviadafgidktacklvavk 866
QY 880	MIKEGATASEHRALMSLEKILIHIGHNLVNVLLGACTKPOGLMIVIVEFCYGNLSNFL 939
DB 867	mikegathsehralmseklilhihghlnvnlilgactkpggplmvivefskfnlstyl 936
QY 940	RAKEDAFSPCAEKSPEQRGRFRA----WVELARLDRRRPGSSDRVLFAFSTEGGARRA 995
DB 927	rkarnefvpyksg-----arfrgkdvygels-vdlikr--rldstssqssasgfvcek 979
QY 996	S-----PDQEAEDLWLSPLTLMEDLVCSYFQVARGMEFLASRKCIHRDLAARNITLSES 1050
DB 980	sldjveeeaseelykdfiltlehlcyfqaakgmeflasrkcihrdlaarnillseknv 1039
QY 1051	VKICDEGLARDIYKDPDYVRKGSARLPKWMAPESIFDKVYTTQSDVWSFGVLLMEIFSL 1110
DB 1040	vkicdfglardiypdyvrkgdaripkwmapietifrvytiqsdvwsfgvllweifsl 1099
QY 1111	GASPYGVOINEEFCCORLRDGTBRAPELATPAIRIMLNCWSGDPKAPASSELVEILG 1170
DB 1100	gaspypgvkideefcrlkegrtmrapdyttemyqtmldcwedpnrqpsfseivehig 1159
QY 1171	DLLQCRGLQEEEEVNCAPRS-SQSSSEGSFVSOTVTHALIAQADARDSPSLQRHSLAAR 1229
DB 1160	nllqanaqdgkdyvlpmsetlsmeedsglsiptspvscmeeeevcdp-----kfh 1211
QY 1230	YVNWVSFPCLARGAETRCSSRMKTFEFPFM-TPPTYKGSVDNQDTSQGMVLASEEPQIE 1288
DB 1212	yntagishylqnskrksrpvsktfedipleevkvipddsqtdgmvlaseelktle 1271
QY 1289	SRHROESGSCKPGQONVAVTRAHPDSQGRRRRPERGARGGVFNSE 1336
DB 1272	drnkispsfg-----gmppsksrsvasegsnq-----tsqygsyghsd 1310
RESULT	8
ID	R37504
AC	R37504 standard; Protein: 1367 AA.
XX	R37504;
XX	19-OCT-1993 (first entry)
DE	Murine fik-1.
KW	Murine; receptor; protein; tyrosine kinase; ptk; fik-1; fik-2; adult;
KW	family; primitive; hematopoietic cell; mature; fetal; liver; spleen;
KW	thymus; brain; stomach; kidney; lung; heart; intestine; bone marrow;
XX	muscle; lymph node; ss
OS	Mus musculus.
XX	Key
XX	Location/Qualifiers
FT	Peptide
FT	/note= "hydrophobic leader sequence"
FT	1..19
FT	/note= "Hydrophobic leader sequence"
FT	20..1367
FT	/note= "Mature fik-1"
FT	20..762
FT	/note= "Extracellular domain"
FT	763..784
FT	/note= "Transmembrane region"
FT	785..1367
FT	/note= "intracellular domain"
PN	W09310136-A.
XX	

Tue Jan 23 12:06:50 2001

PD 27-MAY-1993.
XX 16-NOV-1992; 92WO-US09893.
PF 15-NOV-1991; 91US-0793065.
PR (UYP-R) UNIV PRINCETON.
PA Lemischka IR;
XX WPI; 1993-182479/22.
XX P-PSDB; R37504.
DR Totipotent haematopoietic stem cell receptors, their ligands and
PT DNA sequences - for treating anaemia(s) and bone marrow damage
PT due to e.g. cancer chemotherapy or radiotherapy
XX Claim 40; Fig 2; 127pp; English.
XX This sequence represents the murine receptor protein tyrosine kinase
CC (pTK), flk-1. This pTK is not in the same family as flk-2 (see also
CC Q40914-15) as it is expressed in primitive hematopoietic cells and
CC also in mature hematopoietic cells. flk-1 is expressed in fetal
CC liver, spleen, thymus, brain, stomach, kidney, lung, heart and
CC intestine and adult brain, bone marrow, kidney, heart, spleen, lung,
CC muscle and lymph nodes.
XX Sequence 1367 AA;
SQ
Query Match 37.0%; Score 2665; DB 14; Length 1367;
Best Local Similarity 42.9%; Pred. No. 2.6e-186;
Matches 587; Conservative 205; Mismatches 486; Indels 90; Gaps 24;
QY 1 MORGALCLRLWICLGLDGLVS--GYSWTPPTLNTTEESHVDTGDSLSICRGHPL 58
DB 1 mesgallavalwfcvetraasvpgdflhpkistqkdltilantlqtcgrdld 60
QY 59 WAPCAQAEAPATGDKDSBDTCWRDCECTOARPYCKVLLLEHVNANDGTSGYCYKYKA 118
DB 61 WLPnaq-----rdseervlvecggds-ifckltitprvngndtgaykcsyrdv-- 110
QY 119 RIEGTAASSYVVRDFEOPINPKDT-----LLVNRKDMWVPCLVSLPGLNVL--R 170
DB 111 ----diastvyyvrdyspfiasvsdqhgivytengkctvlpgrgslnlnvslcar 166
QY 171 SOSVLPWPQGEVWMDRRGMLVSTPLLDALYLQCTETTWGDDFLSNPFLVHTGNELY 230
DB 167 ypekrfvpdgnriswdselgflipsymisyagmvfceakindetyqslmyivvvvgyry 226
QY 231 DTQLLPKSLLELVGKLVNLTWMAEENSGVTFDWDYPGKAERGKWWPERSQOHTTE 290
DB 227 dvlsphelelsageklvlnctartelnvgldftcwhspsskshhkvlnrdvxfp9tv 286
QY 291 LS---SILTHNVSHDLGSYVCKANNGFORFESTEVIVHENPFISEWLKGPILEATA 347
DB 287 aknflstltiesvksdqgeytcvassgmikrntfrvhtkpfiafgsgmksliveatv 346
QY 348 GDELVLKLPVLAAYPPPEFOWYKDGKALSGRHS----PHALVLKEVTEASTGYTLALWNS 404
DB 347 gsq-vripkylsyppadikwyrngripiesnytmivgdeltimevterdagnytvlln 405
QY 405 AAGLRNRLSLELVNVVPROTHEKEASSP-SIYSRHSRALCTAYGVPLPLSIOWHWRPW 463
DB 406 ismekqshmslvvnpvppq1gekallispmdsyqgmqlctctvyanpplhhiqwywle 465
QY 464 TPCKMEARSLRRQQDLMPOCRDWRVAVTTQDAVNPIESLDTWTFEVEGKNKTVSKLVI 523
DB 466 eacsy-----rpqgtspackewrhvdfqgggnkievtnqalliegknktvtstivi 517
QY 524 QNANVSAMKCVKSVNKGODERLITYFVTTIPDGFTIESKPSSEELLEGGOPVLLSCQADSY 583
DB 518 qaanvsalykealnagrgervisfhvirgpe---itvqpaaqpteqesvallctadrn 574

QY 584 KYEHLKWRRLNLTSLHDAHGNPLLLDCKNV-----HLFATPLAASLEEVAPGARHATLSLS 639
DB 575 tfenlt-yklsgatsvbmgesltpvcknldalwkngmfmsnstndi-----liva 626
QY 640 IPKVAPEHEGHVCEVQDRSRSHDKHCHKKYLSVOALEAPRLTONLTDLNVNDSLEMOG 699
DB 627 fqnasiqdgdyvcsaqdkktrhclvklillemmapmitgnlenqttitgetietv 686
QY 700 LVACAHAPSIVWYKDERLLEKSGVDLADSNQKLSIQRVREEDAGRYLCSVCNAKCVNS 759
DB 687 pasgnplphitwfkdneltvedsgivldgnrnltrrvrkdggltytqcacnvlcara 746
QY 760 SASVAVEGSEBDKSMIEVILVGTGVIATVFWVWLLLLLIFCNMRPAHADTKTYSIIMDP 819
DB 747 etlfiiegaqektnleivilvgtaviamffwllviltvkranegeiktyglisvmdp 806
QY 820 GEVPLEEQCEVLSYDASQWEPFRERLHLGRVLGYGARGKVVEASAFGIHKGSSCDTVAVK 879
DB 807 delpldrcerlpydaskwefprdrklglpgrgaqgqvieadafgldktatcktvavk 866
QY 880 MKECATASERRALMSELKILIHGNHNLVNVNLLGACTKPGPLWVIVFCVGNLSNLF 939
DB 867 mkegathsehraimseleklilhihgnlnvnlvgactkpgplmiviveskfgnlsyl 926
QY 940 RAKRDAFSPCAEKSPQGRFRER-----MVELARLDRRRPGSSDRVLFARFSTKEGGARRA 995
DB 927 rgkrnefvykskg-----arfrqgkdyvgels-volkr--rldstssqssasgfvcek 979
QY 996 S-----PQGEAEDLWLSPLTMEDLVCSYFQVARGHEFLASRKCITHRLAARNILLSSEDV 1050
DB 980 slsdveeeaseeelykdflltlelicysfvakgmeflasrkcithrldlaarnillseknv 1039
QY 1051 VKICDFGLARDIYKDPDYVRGKSARLPLKWNAPESIFDKVVTQSDVMSFCVLLWEIFSL 1110
DB 1040 vkicdfglardiypdyvrkgdarlplkwnapetifrvytlqsdvswsfvllwelfsl 1099
QY 1111 GASPYGVQVINEFCQRLDGTMRAPELATPAIRIRMLNCWSDGDPKARPAFSELVEILG 1170
DB 1100 gaspygvkideefcrrikegrmradytptemytmdcwhedpnqrpsfseivehlg 1159
QY 1171 DLLOQRGLQEREVEVCMAPRS-SQSEEGSFQSVSTMALHIAQADAEDSPSLQRHSLAAR 1229
DB 1160 nllqanaqdgkdyivlpmssetlsmeedsglsiptspvscmeeevecdp-----kfh 1211
QY 1230 YYNWVSFPCLAGAETRGSSRMKTFEPPM-TPTTYKGSVDNODTSGHVLASEEFDQIE 1288
DB 1212 ydntaqishylqnsk-ksrpsvsktftedipleevkviptdsgtdsgmvlaseelktle 1271
QY 1289 SRHROESGFCCKPGPGONAVTRAHPDSQGRRRRPERGARGGVFNSE 1336
DB 1272 drnklspsfg-----gmmpsksrsvasesgnq-----tsyqsgsyhsd 1310
RESULT 9
R67537
ID R67537 standard; Protein; 1367 AA.
XX
AC R67537;
XX
DT 04-JUL-1995 (first entry)
XX
DE Mouse flk-1.
XX
DE Fetal liver kinase-1; flk-1; protein tyrosine-kinase receptor;
KW hematopoiesis; stem cell.
XX
OS Mus sp.
XX
XX Key 1..19 location/Qualifiers
FT Peptide /label= Sig_peptide
FT Domain 20..762


```

FT      Region      /label= Extracellular_receptor_domain
FT      FT          763..784
FT      Domain      /label= Transmembrane_region
FT      FT          785..1367
FT      FT          /label= Intracellular_catalytic_domain
XX      US5367057-A.
XX      XX
XX      PD      22-NOV-1994.
XX      XX
XX      PF      02-APR-1991; 91US-0679666.
XX      XX
XX      PR      02-APR-1991; 91US-0679666.
XX      PR      28-JUN-1991; 91US-0728913.
XX      PR      15-NOV-1991; 91US-0793065.
XX      PR      24-DEC-1991; 91US-0813593.
XX      PR      26-JUN-1992; 92US-0906397.
XX      PR      12-NOV-1992; 92US-0975049.
XX      PR      19-NOV-1992; 92US-0977451.
XX      PR      30-APR-1993; 93US-0055269.
XX      PA      (UYPR-) UNIV PRINCETON.
XX      PI      Lemischka IR;
XX      XX
XX      DR      WPI; 1995-005894/01.
XX      DR      N-PSDB; Q79070.
XX      XX
XX      PT      Murine flk-2 receptor protein tyrosine kinase - used to stimulate
XX      PT      proliferation and/or stimulation of primitive mammalian
XX      PT      haematopoietic stem cells in vitro or in vivo.
XX      XX
XX      PS      Disclosure; Fig. 3A-3I; 69pp; English.
XX      CC      cdnas encoding receptor protein tyrosine-kinases, mouse fetal liver
XX      CC      kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in Q79068-70.
XX      CC      respectively, and the deduced amino acid sequences in R67535-37,
XX      CC      respectively.
XX      XX
XX      SQ      Sequence 1367 AA;
XX      XX
XX      Query Match      37.0%; Score 2665; DB 16; Length 1367;
XX      Best Local Similarity 42.9%; Pred. No. 2.6e-186;
XX      Matches 587; Conservative 205; Mismatches 486; Indels 90; Gaps 24;
XX      XX
Qy      1 MORGALCLRLWLCGLLDGLVS--GYSMTPPTLNITEESHVIDTQDLSISCRQHPL 58
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      1 mesgallavalwfcvetraasvgipgdfihppkistqkdiilantltlqitcrgqrld 60
Qy      59 WAPCAQAPATGDKDSEDGTGVRCCEGTARPCKVLLHLEHVIANDTGSVCYKVIKA 118
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      61 WLPnaq-----rldeervlvtecgggds-iffctltiprvvngndgagkcsydrdv-- 110
Qy      119 RIETTAASSYVFVRDEFPQPIKPD-----LLVNRKQAMWVPCLVSPGLNVTL--R 170
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      111 ----diastvyvdyrsfpfasvqhgivytkenkntvlpicrgsinlnvsicar 166
Qy      171 SOSVLMFDGQEVVDRRMLVSTPLLDALYLQCTTWCQDQFLNPFVLHITGHELY 230
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      167 ypekrfvpdgnrnsdseigftlpsymisyagmvfceaikndetyqsimyvrvvgvryi 226
Qy      231 DIQLLPKRSLELLYKGLKLVNTWAEFNSGVTFDWDYPCQKQERGWVPERRSQOHT 290
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      227 dvlisppheilsageklvinctartelnvgldftwshpskshhklivardkvpfgvt 286
Qy      291 LS----SILTIHNVSQHDLSGVYCKANNGIQRFRESTEVIVHENPFISVEMWLGPILEATA 347
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      287 akmfistltiesvtskdqgeytcvassgrmlkrrtfrvhtkpfiafsgmkslveatv 346
Qy      348 GDELVKLPVKLAAYPPPEFQWYKDGKALSGRHS---PHALVLRKEVTEASTGTYTLALWNS 404
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      347 'gsk-vripvkylsyapadikwyrrngriplesnytmivgdeltimvetrdagnytviltnp 405

```

```

Qy      405 AAGLRNRISLELVVNVPPQITHEKASSP-SIYSRHSQALCTAYGVLPLSLIQHWRPW 463
Db      : : : : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      406 ismekqshmvslvvn:ppqigekallispmdsyqytmqtltctvyanpphhllqvywql 465
Qy      464 TPKMFAQRSLRRRQODLMPQCRDMRAVTTQDAVNPDESLDTWTEFEVCKNKTVKSLVI 533
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      466 eacsy-----rpgqtspyackewrhvedfggnkievtknqyaliegnktvstlvi 517
Qy      524 QNANVSAMYKCVVSNKVGODERLIYFVTTIPDGFTIESKPSSEELLEGGPVLLSCQADSY 583
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      518 qanvsalykceainkagrgervishvirgpe---itvqaagpteqesvslitcadrn 574
Qy      584 KYEHLRWYRLNLSTLHDAHGNPLLLDCKNV----HLFATPLAASLEEVAFGARHATLSLS 639
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      575 tfenltwyklgsqatsvbmgesltcvcknldalwklingtmfnsnstndi-----liva 626
Qy      640 IPRVAPHEHGYHCEVQDRSHDKHCHKYLSVOALEAPRLTONLTOLLVNVSDSLEMQC 699
Db      : : : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      627 fqnasiqdgdyvcsagdkktrhclvklqilitermapmitgnlenqntttigetievtc 686
Qy      700 LVAGAHAPSIWYKOEKLLLEKSGVDLADSNOKLSIORVEEDAGRYLCSVCNAKGCVNS 759
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      687 pasgnptphitwfkdneltvedsgivrdgnrnltrrvrkedggytqcacnvlgara 746
Qy      760 SASVAVEGSEDKGSMELIVILVGTGVIIVFVFWLLLLIFCNMRRPAHADIKTYLSITMDP 819
Db      : : : : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      747 etlfiiegaqektnlevillvgtaviamfwlllvilvrtvkranegeklgtyslvmdp 806
Qy      820 GEVPLEQCEYLSYDASOMEFPRERLHLGRVLCYGAFCKVVEASAFCIHKSGSCDTVAVK 879
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      807 delpldercerlpydaskwefprdrklgkplgrgavgviedafgtdactcktvavk 866
Qy      880 MLKEGATASEHRALMSELKILIHGNHLNVNVLIGACTKPGGLWTVFECKYGNLSNEL 939
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      867 mlkegathsehalmseklilhighlnvnlilgactkpggplmvlviefskgnlstyl 936
Qy      940 RAKRDAPSPCAEKSPQGRFRFA----WVELARLDRRRPGSSDRVLFARFSTEGGARRA 995
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      927 rgkrnefvykyskg---arfrgqkdyvgels-vdlkr--rldsitssqssasgfvcek 979
Qy      996 S-----PDQEAEDLWLSPLTMDLVCSFOVARGMEFLASRKCITHRDLAARNILLSSESV 1050
Db      : : : : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      980 slsdveeeaseelykdfitclehlicysfqvkgmeflasrkcihrdlaarnillseknv 1039
Qy      1051 VKICDFGLARDIYKDPDYVRKGSARLPKWMAPESIFDKYVTTQSDVMSFGVILLWEIFSL 1110
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      1040 vkicdfglardiypdyvrkgdarlpkwmapietifdrvytiqsdvmsfgvillweifs 1099
Qy      1111 GASPYPCVQINEEFCQHLRGTWRAPELATPAIRRIIMLCWSCDQKARPAFSELVEILG 1170
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      1100 gaspyppvkiideefcrlkegtrmrappyttemyqlmldcwhehdpnqrpsfseivehlg 1159
Qy      1171 DLLQGRLOQEEVEVCNAPRS-SQSEGESFSQVSTMAHIAQADAEPSPLQRHSLAAR 1229
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      1160 nllqanaqqdgkyivlpmssetlsmeedsglsiptspvsmeeveevcdp-----kfh 1211
Qy      1230 YHNVSPFGCLARGAETRGSSRMKTFEFPN-TPTTYKGSVDNOTSGMYLASEEFOIE 1288
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      1212 ydtagishylqnskrsvsvktfedipleepevkvipddsgtdsgmvlaseeiktle 1271
Qy      1289 SRHROESGESCCKPGQNVAVTRAHPDQSGRRRRPERCARGQOVFYANSE 1336
Db      I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db      1272 drnklspsfg----gmmpsksvsasegsnq-----tsgyqsgyghsd 1310

```

```

RESULT 10
R97420
ID R97420 standard; Protein; 1367 AA.
XX
AC R97420:
XX
DT 11-DEC-1996 (first entry)

```


form of flk-2 which is used to isolate specific ligands for flk-2. These ligands can be used to stimulate proliferation and/or differentiation of mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for treatment of macrocytic or aplastic anaemia or bone marrow damage caused by cancer treatment or radiation.

Sequence 1367 AA;

Query Match 37.0%; Score 2665; DB 18; Length 1367;

Best Local Similarity 42.9%; Pred. No. 2.6e-186;

Matches 587; Conservative 205; Mismatches 486; Indels 90; Gaps 24;

QY 1 MORGAAALCLRLWLCGLDGLVLS--GYSMTPTPLNITEESHVIDTGDLSISCRGQHLE 58

DB 1 meskgllavalwlcetraasvlgldfhhpkalslqkdltilantllqicrgrld 60

QY 59 WAMPGAOEAPATGDKDSEDTGVVRDCEGTDAEPYCKVLLHEVHANDTGSVCYKYKA 118

DB 61 wlpnaq-----rdseervlvtecgsgds-ikfktltprrvngndgkaycsyrdv-- 110

QY 119 RIEGTTAASSYVFEVFEQPFINKPDT-----LLVNRKDAWVPCVLSIPGLNVTL--R 170

DB 111 ----diastvvyvdyrpsfiasvdsdqhgivitenkktvviptgsgsinlvsicar 166

QY 171 SOSVLPDQGVVWDDRRGMLVSTPLLDALYLQCTTWTGDDQDFLSNPLVHITNELY 230

DB 167 yekrfvpgdnriswdsiegtfipsymisyagmvceakindetyqsimyivvvgyr 226

QY 231 DIQLLPKRLLELVGEKLVINCTVWAEFNSGVTFDWDYPCOAERKWPERRSQOHTTE 290

DB 227 dvlspheiselsageklvinctartelnvgldftwhspbskshhkhivndrvkpfpgtv 286

QY 291 LS----SILTHNVSOHDLGSVVCCKANQIORFRESTEVIHVENPFIIVENLKGPILEATA 347

DB 287 akmfstltitvstskdggcytcvassgrmiknrtfvrvtkpfafgsgmkslveatv 346

QY 348 GDELVKLPVLAAYPPPEFQWYKDGKALSGRHS----PHALVLKEVTEASTGTTLALWNS 404

DB 347 gsq-vripvkylsyapdikwyrngriplesnytmilvgdeltmevterdagnylvltnp 405

QY 405 AAGLRNLSLELVNVPQIHEKEASSP-SIYSRHSQALCTATAYVPLPLSIQHRNPW 463

DB 406 ismekqshmvslvnnvpqigekalispmsdyqygtmqtlctctvyanpplhhqvyqle 465

QY 464 TPKMFAQRSLRRQOQDLNFCQDRWRAVTTODAVNPVIESLDTWTEFEVGRKNTVSKLVI 523

DB 466 eacsy-----rpgqlspyakewrhvedfgggnklevtknqyallegknktvstivi 517

QY 524 QNANVSAMYKCVVSNKVGQDERLIYFVYVTPDGTFTIESKPSEELLEGOPVLLSQADSY 583

DB 518 qanvsalykceainkagrgervisfhlrgpe---ilvqpaadqteqesvllctadrn 574

QY 584 KYEHLRWYRLNLSTLHDAHGNPLLDCKNV-----HLFATPLAASLEEVAPGARHATLSLS 639

DB 575 tlenitvylgsgatsvbmgesltpvcknldalwklnglufmfnstndi-----liva 626

QY 640 IPKVAPEHGHVYCEVQDRRSHDKHKYLSVQALEAPRLQNTLTLVNVSDSLEMQC 699

DB 627 fgnasldqddgyvcsaqdkktrhclvklillemepmitgnlenqtttietietvc 686

QY 700 LVAGHAPSIVMYKDERLLEESGVLDLADSNOKLSIQRVREEDAGRYLCSVSNACKVCVNS 759

DB 687 pasgnptphitwfkdneltvedsgivlrdgnrnltrrvkedgglytcqacnvgcara 746

QY 760 SASVAVEGSEDKSMEIVLVGTCVIAVFFWVLLILLIFCNMRPRAHADIKTGVLSIIMP 819

DB 747 etifiieaqeknlvliivgtaviamffvllilvtrvkranegeklgtglisvmdp 806

QY 820 GEVPLEQCEYLVSDASOWEFPFRERLHGLVGLYGFAGKVFVEASAFGIHKSSCDVAVK 879

DB 807 delpidercerlpydaskewefdrklgkplgrgafgvgviedafgidktatctvavk 866

DB 1100 gasyppgvkideefrrlkegtmrpadtyttemyqtmldcwshednqrpsfvelhlg 1159

QY 1171 DLQAGRGLOEEVVCNAPRS-SOSSEGSFQSVTWMALHIAQADAEDSPSLORHSLAAR 1229

DB 1160 nilqanaqqgdkyivlpmetismedsglsiptspvscmeeeevcdp-----kfh 1211

QY 1230 YNNVSPFGCLARCAETRGSSRMKTEEFPM-TPTTYKGSVDNQTSQWVLAEEFEQIE 1288

DB 1212 ydntagishylqnskrkrpsvsktfdipleepekvipddsqtdsgmvlaseelktle 1271

QY 1289 SRHQESFCFGQGVNVAHVTRAHPDSQGRRRRPERGARGGQGVFNSE 1336

DB 1272 drnklspsfg----gmmpsksrsvasegnsq-----tsyqsgyhsd 1310

RESULT 11

ID W19875 standard; Protein; 1367 AA.

AC W19875;

DT 19-AUG-1997 (first entry)

DE Murine flk-1 receptor.

Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase; ptk; liver; spleen; thymus; adult; brain; bone marrow; primitive portion; haematopoietic hierarchy; extracellular domain; soluble form; ligand; proliferation; differentiation; mammalian; haematopoietic stem cell; macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.

OS Mus musculus.

FH Key Location/Qualifiers

FT Region 763..784

FT /note= "Transmembrane region"

XX US5621090-A.

XX 15-APR-1997.

XX 02-APR-1991; 91US-0679666.

XX 26-JUN-1992; 92US-0906397.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPT; 1997-235228/21.

XX N-PSDB; 772119.

XX Protein containing the extracellular domain of human flk-2 - used for identification of primitive haematopoietic cell proliferation and differentiation stimulatory ligands, e.g. for treating anaemia

XX Disclosure; Fig 2; 55pp; English.

XX This sequence represents murine fetal liver kinase 1 (flk-1). flk-1 is a receptor protein tyrosine kinase (ptk). flk-1 is found in the same organs as flk-2 (see also 72117-18) as well as in fetal brain, stomach, kidney, lung, heart and intestine, and in adult kidney, heart, spleen, lung, muscle and lymph nodes. flk-1 is not a member of the same class of receptors as flk-2, as flk-1 may be found in more mature haematopoietic cells. The invention concerns a recombinant nucleic acid, preferably mRNA, which encodes a protein containing only the extracellular domain of human flk-2 and lacking the flk-2 intracellular catalytic domain. The resultant protein represents a soluble

QY 880 MLKAGATASEHRAIMSELKILIHGHNHNVNLLGACTKPKQGLMIVIFCKYKGNLSNFL 939
 Db 867 mlkegathsehraimseikililighlnvvnllgactkpggplmivieskfgnlstyl 926
 QY 940 RAKRDAFSPCAEKSPQGRPR-----WVELARLDRRPGSSDRVLPARSKTEGGARRA 995
 Db 927 rgknefvyyskg-----arfrqgdyvgels-vdlkr--rldsitssqasagfveek 979
 QY 996 S-----PDQEAEDLWLSPLTMEDLVCSFOVARGMEFLASRKCIHRLAARNILLSSDV 1050
 Db 980 slsveeeaseelykdfitlehlicysfqvqgmeflasrkchrlaarnillseknv 1039
 QY 1051 VKICDFGLARDIYKDDPVYRKSGARLPLKWMAPESIFDKVYTTQSOWMSFGVLLWEIFSL 1110
 Db 1040 vkicdfglardiykddpvyrkgdarlpkwmapietifdrvytiqsdvswsfvllweifs 1099
 QY 1111 GASPPYGVQINDEFQORLDRCTMRAPLATPATIRITMLNCWSDGPKARPAFSELVEILG 1170
 Db 1100 gaspypgvkideefcrrikegtrmradytptemyqlmcdwhedpqrpsfiselvehlg 1159
 QY 1171 DLQGRGLQEEVEVCMAPRS--SQSSEGSFQSVSTMALHIAQADAEDSPPSLQRHSLAAR 1229
 Db 1160 nilqanaqdgkdyivlpmssetlsmeedgslsptpsvscmeeevcdp-----kfh 1211
 QY 1230 YNNWVSFPGCLARGAETRGSRMKTFFEFPM--TPTTYKGSVDNQTDSGMVLASEEFOIE 1288
 Db 1212 yntagishyiqnskrprsvsktfedipleepekvipdsqtdsgmvlaseelktle 1271
 QY 1289 SRHROESGFCGKQCONVAVTRAHPDSQGRRRRPERGARGQVFNSE 1336
 Db 1272 drnklspsfg---gmmpskresvasegnsq-----tsyqsgyhsd 1310

RESULT 12

Y08618
 ID Y08618 standard; Protein; 1367 AA.

XX AC Y08618;

XX DT 05-AUG-1999 (first entry)

XX DE Murine flk-1 protein.

XX KW Murine; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;
 monoclinal; polyclonal; antibody; tyrosine kinase.

XX OS Mus sp.

XX PN US5912133-A.

XX PD 15-JUN-1999.

XX PF 10-FEB-1998; 98US-0021324.

XX PR 19-NOV-1992; 92US-0977451.

XX PR 02-APR-1991; 91US-0679666.

XX PR 28-JUN-1991; 91US-0728913.

XX PR 15-NOV-1991; 91US-0793065.

XX PR 24-DEC-1991; 91US-0813593.

XX PR 26-JUN-1992; 92US-0906397.

XX PR 12-NOV-1992; 92US-0975049.

XX PR 30-APR-1993; 93US-0055269.

XX PR 31-OCT-1994; 94US-0252498.

XX PR 15-FEB-1996; 96US-0601891.

XX PA (UYPR-) UNIV PRINCETON.

XX PI Lemischka IR;

XX DR WPI; 1999-357194/30.

XX DR N-PSDB; X77516.

XX PT Isolating hematopoietic cells expressing fetal liver kinase 1

PT receptors
 XX
 PS
 CC Disclosure; Fig 2; 59pp; English.
 CC This invention describes a novel method of isolating cells expressing fetal liver kinase 1 (flk-1) receptors on their surface and comprises binding the cells to a polyclonal or monoclonal antibody specific to the flk-1 receptor and isolating the cells that have bound to the antibody. The method can be used to isolate hematopoietic stem cells in any mammal but preferably a rat, mouse, rabbit or human. The proteins of the invention belong to the receptor protein family. This sequence represents the murine flk-1 protein which is used in the method of the invention.

XX
 SQ Sequence 1367 AA;

Query Match 37.0%; Score 2665; DB 20; Length 1367;
 Best Local Similarity 42.9%; Pred. No. 2,6e-186;
 Matches 587; Conservative 205; Mismatches 486; Indels 90; Gaps 24;

QY 1 MORGAAALCLRLCLGLDGLVS--GYSMPTPLNITEESHVIDTGSLSISCRGQHPLE 58
 Db 1 meskgllavalwfcvetraasvlgpdlfnpklskqdltilantlqitcrgardid 60
 QY 59 WAWPGAOEAPATGDKDSEDGTGVVRDCEGTDAIPYCKVLLHEVHANDTGSVYCYKIKA 118
 Db 61 wlpwnag-----rdseervlvtecgggds-iffckltltpvrvvgnatgkcsydv-- 110
 QY 119 RIEGTTAASSYVFVRDPEOFINKPDT-----LLVNRKDAMWVCLVSLIPGLNVTL--R 170
 Db 111 ----diastvvyvdyrpsflasvdsqhgivytkenkntkvtvpcrgslnlnsvicar 166
 QY 171 SQSSVLWPDQOEYVWDDRRGHLVSTPLLDHALYLOQETGWDGDFLSNPFVHITGNELY 230
 Db 167 ypekrtpvdpnriswdseigftlpsymisvaygmvfceakindetyqslmyivvvgyriy 226
 QY 231 DICLLPRKSLLELLVGEKLVNCTVWAEFNSGVTFDWDYFGKQAEKRWVPERSSQOHTHE 290
 Db 227 dvilsphelelsageklvinctartelnvgldftvhsppskskhkknvrdvdkpfgtv 286
 QY 291 LS---SILTIHNVSOHDLGYSVCKANNQORFRESTEVIVHENPFISVEWLKGPILEATA 347
 Db 287 axmifstultiesvtksdqgeytcvassgmiknrtfvrvtkpfiafsgmsklsveatv 346
 QY 348 GDELVKLPVKLAAYPPPEFQWYKDGKALSGRHS---PHALVLKVEYEAATGYVTLALMS 404
 Db 347 gsq-vripvkylspapdikwyrngripiesnytmivgdeltmevterdagnvtviltap 405
 QY 405 AAGLRNLSLELVNVVPPQIHEKEASSP-SIYSRHSRQALCTCTAYGVPLSLTQWHRPW 463
 Db 406 ismekgshmvslvnnvppqlgekalispndsyqgtmqtlctctvyanpplhhiqwyqlie 465
 QY 464 TPCKMFAQRSLRRRQOQDLMPQCRDMRAVTTQDAVNPIESLDTWTVEFGKKNKTVSKLVI 523
 Db 466 eacsy-----rpqstpyackewrhvedfqqgnkievtknqyaliegnkktvstlvi 517
 QY 524 QNANVSAMYKCVVSNKVGQDERLIYFYVTTIPDGFTIESKPSSEELLEGGPVLSCQADSY 583
 Db 518 qaanvsalykceainkagrgervisfhvirgpe---itvqaaqpteqesvsilctadn 574
 QY 584 KYEHLRWYRLNLSTLHDAHGNPLLLDCKNV---HLFATPLAASLEVEAPGARHATLSLS 639
 Db 575 tfenltwyklsgqatsvhngesltvpycknldalwkngcmfsnstndi-----liwa 626
 QY 640 IPKVAPEHGHVYCEYQDRRSHDKCHKKYSVQALEAPRLTQNLTDLLVNVSDSLEMQC 699
 Db 627 fqnasiqdgdyvcsaqdktkkrhclvklqllilermapiitgnlenqntttigetievlc 686
 QY 700 LVAGAHAPSIVWYKDERLLEKSGVDLADSNOKLSIORVEEDAGRYLCSVCAKCCVNS 759
 Db 687 pasgnplnltwfkdnethivedsglvirdgnrntirrvrkedgglytqcacnvlgcara 746

Qy	760	SASVAVGSESDKSGMSIEIVILGTGVIAVFFWVLLLLLIFCNMRRRAHADIKTGYLSLIIMDP	819
Db	747	etlfiieagakeknleivllgtavilamfllwllvrtvkraneelktgylsivmvp	806
Qy	820	GEVPLEQCEVLSDYASOWEPPERLHLGRVLGVCAGFCVVEASAFGIHKSGSCDTAVAK	879
Db	807	deipldercerlpydaskwekprdrlikgplgrgafgqvleadafgldtatcktvavk	866
Qy	880	MLKEGATASEHRALMSLKLIIHIGNHNVNLLGACTKPOGQLMWIVECKYGNLSNFL	939
Db	867	mlkegathsehralmseklilhihghlnvvnllgactkpgplmwivefskfgnlstyl	926
Qy	940	RAKRDASFPCAESKPEQRGRFA---MVLEARLDRRRPGSSDRVLFAFSTKEGGARRA	995
Db	927	rgkrnetfpyksg---arfggkdygels-vdlkr--rldsitssqssasgfvweek	979
Qy	996	S-----PDQEAEDLWLSPTMEDLVCYSFOVARGMEFLASRKCITHRDLAARNILLSVDY	1050
Db	980	slsdveeeaseelykdfitlehlicysfvakgmeflaskchrdlaarnillseknv	1039
Qy	1051	VKICDFGLARDIYKDPYVRKSGARLPLKWMAPESIFDKVYTTQSDVMSFGVLLWEIFSL	1110
Db	1040	vkicdfglardiypdyvrkgdarlplkwmapeifdrvvtlqsdvwsfgvllweifsl	1099
Qy	1111	GASPYGVOINEEFCORLDRGTMRAPELATPAIRIMLNCWSDGPKARPAFSELVEILG	1170
Db	1100	gaspygvkideefrcrlkegtrmapdyitpemyqtmldcwhebpqrfsfsevehlg	1159
Qy	1171	DLQGRGLOEEVEECVMAKPS-SOSSEEGSFQSVSTMALHIAQADAEDSPPSLQRHSLAAR	1229
Db	1160	nllqanaqqgkdyivlpmsettlsmeedsglslpsvscmeeeevcdp-----kfh	1211
Qy	1230	YNNWVSFGCLARGAETRGSSRNKTFEEFPM-TPTYKGSVDNOTDGGVLAASEEFOIE	1288
Db	1212	ydntagishylqnskrrkspvsktfdipleepekvipddsqldsgmvlaseelktle	1271
Qy	1289	SRHROESFGCKPGQNVATRAHPDSQCRRRRRPERGARGQGVFNSE	1336
Db	1272	drnklspsfg---gmmpsksrsvasegsnq-----tsgyqsgyhsd	1310
RESULT 13			
R28041			
ID	R28041 standard; Protein; 1367 AA.		
XX	R28041;		
AC			
XX	15-MAR-1993 (first entry)		
Df			
XX	flk-1.		
DE			
XX			
KW	Thymidine kinase; TK; haematopoietic; stem cells; proliferation;		
KW	differentiation; progenitor cells; foetal liver kinase.		
XX			
OS	Mus musculus.		
XX			
FH	Location/Qualifiers		
FT	763..784		
FT	/note= "transmembrane region"		
XX			
PN	W09217486-A.		
XX			
PD	15-OCT-1992.		
XX			
PF	02-APR-1992; 92MO-US02750.		
XX			
PR	02-APR-1991; 91US-0679666.		
PR	28-JUN-1991; 91US-0728913.		
PR	15-NOV-1991; 91US-0793065.		
PR	24-DEC-1991; 91US-0813593.		
XX			
XX	(UYPR-) UNIV PRINCETON.		
XX			

QY 405 AAGLRNLSLELVVNPPOIHEKEASSP-SYSRHSROALTCTAYCVPLPLSIOWHRPW 463
 Db 406 ismekqshmslvvnpvppdigeakalispmdsyqymqltctvyanplhhiqwywle 465
 QY 464 TPCKMEAQRSLRRROODLMPQCRDWRAYVTTQDAVNPIESLDTWTFVEGKNKTKVLYI 523
 Db 466 eacsy-----rpgqtpyackewrhvedfggnkievtknqvaliegnkktvtivi 517
 QY 524 QNANVSAMTKCVSVNKVGDERLIYFYVTIPDGFTIESKPSSEELLEGQPVLLSCQADSY 583
 Db 518 qanvsalykceainkagrgervishvirgpe---itvqpaqpteqesvsallctadrn 574
 QY 584 KYEHLRWYRLNLSTLHDHGNPLLLDCKNV---HLFATPLAASLEEVAPGARHATLSLS 639
 Db 575 tfenltwykgsqatsvhngesltvpcnkldalwknglmgfnsntndi-----liwa 626
 QY 640 IPRVAPEHEGHVYCEVDRRSHDKHCKKYSVQALEAPRLTONLTDLNVNVSLSLWQOC 699
 Db 627 fqnasiqdgdyvcsaqdkktrhclvklqililermapmitgnlenqntttigetievtc 686
 QY 700 LVAGAHAPSIVWYKDERLLEKSGVDLADSNOKLSTORVEDAGRYLCSVCNAKCVNS 759
 Db 687 pasngptphltwfkdelvedgvlrdgnrnltrvrkedggltycqacnvlgcara 746
 QY 760 SASVAVGSEDKGSMETILVGTGTAFFVWVLLLLIFCNMRPAPAHADIKTYGLSTINDP 819
 Db 747 etlfiegaqektnlevlllvgtaviamfflllvtvkvranegelktgylsvmdp 806
 QY 820 GEVPLESOCEYLSYDASOWEPPRERHLGRVLYGYAGFKVVEASATGIHKGSCDVAVK 879
 Db 807 delpldercerlypdaskwefprdklqkplgrgafgvgviedafgdktdatcktvavk 866
 QY 880 MLKEGATASRRALMSLKLTIHGHNLNVNVLGACTKPOGLWLVFECKYGNLSNLF 939
 Db 867 mlkegasenralmselklllhighnlvnnllgactkpggplmvlfvskfnglntyl 926
 QY 940 RAKDAFSPCAEKSPQGRFRA-----MVELARLDRRRPGSSDRVLFAFASKTEGGARRA 995
 Db 927 rgrknefvpkykg---arfrgqkdyvgels-vdikr---rldsltsqssasgfvcek 979
 QY 996 S-----PDQAEADLWLSPLTMEDLVCSFOVARGMEFLASRKCIHRDLAARNLLSES DV 1050
 Db 980 slsdeveeaeaseelykfltiehlycsfqvqkmgfslarkicldlaarnllseknv 1039
 QY 1051 VKICDFGLARDIYKDPDYVRKGSARPLKWMAPESIFDKYVTTQSDVWMSFGVLLWEIFSL 1110
 Db 1040 vkicdfglardiypdyvrkgsdarplkwmapietifdrvytiqsdvwsfgvllweifsl 1099
 QY 1111 GASPYPGVQINEEPCQRLRDGTRNAPELATPAIRRIIMLNCWSDCPKARPAFSELVEILG 1170
 Db 1100 gaspypgvkideefcrriketgrmradyttpeymqtlmldcwhepnpqrfselvehlg 1159
 QY 1171 DLGGRGLQEEVEECMAPRS-SOSSEEGSFQVSTMALHTAQADAEDSPSLQRHSLAAR 1229
 Db 1160 nllqanqqdgkdyivlpmetismedsglsiptspvcmeeevcdp-----kfh 1211
 QY 1230 YNNWVSFPGCLARGAETRGSSRMKTFEFPM-TPTTYKGSVDNQTOSGMVLASEEFPQIE 1288
 Db 1212 ydntagishylqnskrksrpsvktfedipleepevkvipddsqtdsgmvlaseelktle 1271
 QY 1289 SRHROESGFCGKQCNVAVTRAHPDSQGRRRRRPARGCGOVFNSE 1336
 Db 1272 drnklspsfg-----gmmpsksvresvasegnsq-----tsyqsgyhsd 1310

RESULT 15

R54046

ID R54046 standard; Protein; 1367 AA.

XX AC R54046;

XX AC R54046;

DT 08-NOV-1994 (first entry)

XX

DE Sequence of murine foetal liver kinase 1 (Flk-1).

KW Foetal liver kinase; receptor tyrosine kinase.

XX Mus musculus.

OS WO9411499-A.

PN 26-MAY-1994.

PD 15-NOV-1993; 93WO-EP03191.

PF 13-NOV-1992; 92US-0975750.

PR 26-MAR-1993; 93US-0038596.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA Millauer B, Risau W, Ullrich A;

PI WPI; 1994-183501/22.

DR N-PSDB; Q64049.

XX DNA encoding Flk-1, a tyrosine kinase receptor for vascular endothelial growth factor - used to express recombinant Flk-1 for screening for ligands useful for modulating vasculogenesis and angiogenesis e.g. for treating cancer

PS Disclosure; Page 50-56; 99pp; English.

CC A receptor tyrosine kinase cDNA, designated foetal liver kinase 1 (Flk-1), was cloned from mouse cell populations enriched for haematopoietic stem and progenitor cells. The nucleotide coding sequence and deduced AA sequence of the murine Flk-1 gene is depicted in Q64049/R54046 and has been described in Matthews et al., 1991, Proc. Natl. Acad. Sci. USA, 88:9026-9030. The murine Flk-1 gene was isolated by performing a PCR using two degenerate oligo primer pools that were designed on the basis of highly conserved sequences within the kinase domain of receptor tyrosine kinases (Hanks et al., 1988). Based on AA homology, this receptor is a member of the type III subclass of RTKs (Ullrich and Schlessinger) which could immunoglobulin-like repeats in their extracellular domains.

XX Sequence 1367 AA;

Query Match 36.6%; Score 2642; DB 15; Length 1367;

Best Local Similarity 42.7%; Pred. No. 1.2e-184;

Matches 584; Conservative 206; Mismatches 488; Indels 90; Gaps 24;

QY 1 MORGAAALRLWLICLLDGLV--SCYSMTPTPLNITEESHVIDTGDLSISICRGQHPLE 58

Db 1 meskallavalwfcvetraasvqltgdfhnpklstcqdltilantlqitcrgdrld 60

QY 59 WAMPGAQEPATGDKDSEDTGVVRDCEGTDARPYCKVLLHVEHANDGTSYVCYKYIKA 118

Db 61 clvpnaq-----rdseervlvtecgqds-ikfktltpvrvngndgkaycsyrdv-- 110

QY 119 RIEGTTAASVYVVRQFQPFINKPT-----LLVNRKDMVPCVLSTPGLNVTL--R 170

Db 111 ----diastvyvdyrpsflasvdsqhgvyitenkntkvtvpcrgslsnlvsicar 166

QY 171 SQSSVLWPQGEVWMDRRGMLVSTPLLDALYLQCETTWGDQDFLSNPFLVHITGNELY 230

Db 167 ypekrfpvdpnriswdselgftipsymisyagmvfcaekindetyqsimyivvvgyrly 226

QY 231 DIQLPRKSLELLVGEKLVNCTVWAEFFNSGVVTFDMQYPCQKAEGRKWWPERSQOQTHE 290

Db 227 dvilsphelelsageklvlnctartelnvgldftchspbpskshhkiivrndvkvpgvtv 286

QY 291 LS---SILLTHNVSOHDLGYSVCKKANGGIORFRESTEVIHNPFIIVENLKPFILEATA 347

Db 287 akmfllstltiesvtrksdqgeytcvassgrmknrtfvrvtvkpfiafgsgmkslveatv 346

[illegible]

Sequence 5, Appli
Sequence 3, Appli
Sequence 18, Appli
Sequence 18, Appli
Sequence 5, Appli
Sequence 15, Appli
Sequence 15, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 17, Appli
Sequence 16, Appli
Sequence 36, Appli
Sequence 36, Appli
Sequence 4, Appli
Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-340-011-4
: Sequence 4, Application US/08340011
: Patent No. 5776755
: GENERAL INFORMATION:
: APPLICANT: Alitalo, et al.
: TITLE OF INVENTION: FLT4, A NOVEL RECEPTOR TYROSIN KINASE
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/340.011
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/959,951
: FILING DATE: 09-OCT-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Thomas C.
: REGISTRATION NUMBER: 36,989
: REFERENCE/DOCKET NUMBER: 32267
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: TELEX: 25-3856
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1363 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-340-011-4

Query Match 100.0%; Score 7211; DB 1; Length 1363;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQRGAALCLRLWLCGLGLDGLVSGYMTPTLNITEHSVIDTGDLSISCRQHLEWA 60
DB 1 MQRGAALCLRLWLCGLGLDGLVSGYMTPTLNITEHSVIDTGDLSISCRQHLEWA 60

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: January 23, 2001, 04:55:57 ; Search time 53.54 Seconds
(without alignments)
457.143 Million cell updates/sec

Title: US-09-375-248-2
Perfect score: 7211
Sequence: 1 MQRGAALCLRLWLCGLGLDGLVSGYMTPTLNITEHSVIDTGDLSISCRQHLEWA 1363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues
Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7211	100.0	1363	1	US-08-340-011-4
2	7211	100.0	1363	2	US-08-901-710-5
3	7157	99.3	1363	2	US-08-874-678-32
4	7157	99.3	1363	3	US-08-643-839-32
5	6852	95.0	1298	1	US-08-222-616-33
6	6852	95.0	1298	1	US-08-340-011-2
7	6852	95.0	1298	3	US-08-901-710-2
8	6852	95.0	1298	4	PCT-US95-04228-33
9	6797.5	94.3	1368	2	US-08-874-678-34
10	6797.5	94.3	1368	3	US-08-643-839-34
11	6005.5	83.3	1362	2	US-08-874-678-33
12	6005.5	83.3	1362	3	US-08-643-839-33
13	4135	57.3	777	2	US-08-874-678-3
14	4135	57.3	777	3	US-08-643-839-3
15	2753	38.2	1356	1	US-08-810-116-8
16	2753	38.2	1356	2	US-07-930-548A-8
17	2682	37.2	1367	2	US-08-443-861-2
18	2665	37.0	1367	1	US-07-813-593-4
19	2665	37.0	1367	1	US-07-977-451-6
20	2665	37.0	1367	1	US-07-946-507-4
21	2665	37.0	1367	1	US-08-252-517-6
22	2665	37.0	1367	1	US-07-906-397A-6
23	2665	37.0	1367	1	US-08-601-891-6
24	2665	37.0	1367	2	US-09-021-324-6
25	2665	37.0	1367	4	PCT-US92-02750-8
26	2665	37.0	1367	4	PCT-US92-05401-6
27	2665	37.0	1367	4	PCT-US92-09893-6
28	2501.5	34.7	1311	1	US-08-340-011-5

1141 TPAIRIMLNCWSDPKARPAFSELVEILGDLQGRLOEEVEEVCMPARSSOSSEGSFS 1200
1141 TPAIRIMLNCWSDPKARPAFSELVEILGDLQGRLOEEVEEVCMPARSSOSSEGSFS 1200
1201 QVSTHALHIAQADSDSPSLQORHSLAARYNNWSPFCLARGAETRGSSRMKTFEEFPM 1260
1201 QVSTHALHIAQADSDSPSLQORHSLAARYNNWSPFCLARGAETRGSSRMKTFEEFPM 1260
1261 TPTTKGSDVNDOTSGMVLASEFEQIESRHROESGFCCKPGQNVAVTRAHPDSQGRRR 1320
1261 TPTTKGSDVNDOTSGMVLASEFEQIESRHROESGFCCKPGQNVAVTRAHPDSQGRRR 1320
1321 RPERGARGOVFNSEYSELSEEDHCSPSARVTFETDMSY 1363
1321 RPERGARGOVFNSEYSELSEEDHCSPSARVTFETDMSY 1363

RESULT 2
US-08-901-710-4
: Sequence 4, Application US/08901710
: Patent No. 6107046
: GENERAL INFORMATION:
: APPLICANT: Alitalo, Kari
: APPLICANT: Aprelikova, Olga
: APPLICANT: Pajusola, Katri
: APPLICANT: Armstrong, Elna
: APPLICANT: Korhonen, Jaana
: APPLICANT: Kaipainen, Arja
: APPLICANT: Matikainen, Marja-Terttu
: TITLE OF INVENTION: FLT4, A RECEPTOR TYROSINE KINASE, AND USES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/901,710
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/340,011
: FILING DATE: 14-NOV-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/257,754
: FILING DATE: 09-JUL-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/959,951
: FILING DATE: 09-OCT-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Gass, David A.
: REGISTRATION NUMBER: 38,153
: REFERENCE/DOCKET NUMBER: 28113/33824
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: TELETYPE: 25-3856
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1363 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-901-710-4

61 WPGAQAPATGDKDSEDGTGVRDCEGTDPAPYCKVLLHEVHNDTGSVCYCYKIKARI 120
61 WPGAQAPATGDKDSEDGTGVRDCEGTDPAPYCKVLLHEVHNDTGSVCYCYKIKARI 120
121 EGTAAASSYVFRDQPFINKPDTLLVNRKDMWPCLVSTPGLNVTLSRSSVLPDQ 180
121 EGTAAASSYVFRDQPFINKPDTLLVNRKDMWPCLVSTPGLNVTLSRSSVLPDQ 180
181 QEVWMDRRGMLVSTPLLDALYLQCEETWGDQDFLSNPFVHITGNELDYDQLPRKSL 240
181 QEVWMDRRGMLVSTPLLDALYLQCEETWGDQDFLSNPFVHITGNELDYDQLPRKSL 240
241 ELLVGEKLVNCTVWAEFNSGVTFDNDYDCKOAEKRWPERRSQOHTLSILTIHNV 300
241 ELLVGEKLVNCTVWAEFNSGVTFDNDYDCKOAEKRWPERRSQOHTLSILTIHNV 300
301 SQHDLGSYVCKANNGIORRESTEVIHNPFTISVEMLKGPILDEATAGDELVKLPVLA 360
301 SQHDLGSYVCKANNGIORRESTEVIHNPFTISVEMLKGPILDEATAGDELVKLPVLA 360
361 YPPPEFQWYKDGKALSGRHSFHALVLKEVTEASTGYTTLALNWSAAGLRNLSLELVNV 420
361 YPPPEFQWYKDGKALSGRHSFHALVLKEVTEASTGYTTLALNWSAAGLRNLSLELVNV 420
421 PPQIHEKEASSPSIYSRHSQALCTAYGVPLPLSIQWHRPWTCKMFAQSRSLRRQOQ 480
421 PPQIHEKEASSPSIYSRHSQALCTAYGVPLPLSIQWHRPWTCKMFAQSRSLRRQOQ 480
481 DLMPCQDRMRAVTTQDAVNPISLDTWTEVEGKNKTVSKLVTONANVSAMKCVVSNKV 540
481 DLMPCQDRMRAVTTQDAVNPISLDTWTEVEGKNKTVSKLVTONANVSAMKCVVSNKV 540
541 GQDERLIYVYTTIPDGFTEISPSBELLEGQPVLLSCQADSKYKHEHLRWYRLNLTLD 600
541 GQDERLIYVYTTIPDGFTEISPSBELLEGQPVLLSCQADSKYKHEHLRWYRLNLTLD 600
601 AHGNPLLLDKKNVLFATPLAASLEEVAPGARHATLSLSPRVAPEHEGHVYCEVQDRS 660
601 AHGNPLLLDKKNVLFATPLAASLEEVAPGARHATLSLSPRVAPEHEGHVYCEVQDRS 660
661 HDKICHKKYLSVQALEAPRLTQNTDLLVNVSDLEQVLLSCQADSKYKHEHLRWYRLNLTLD 720
661 HDKICHKKYLSVQALEAPRLTQNTDLLVNVSDLEQVLLSCQADSKYKHEHLRWYRLNLTLD 720
721 KSGVDLADSNOKLSIQVRREDAGRYLCSVCNARGCVNSSASVAVEGSDKGSMEIVILV 780
721 KSGVDLADSNOKLSIQVRREDAGRYLCSVCNARGCVNSSASVAVEGSDKGSMEIVILV 780
781 GTGVIAVFFWVLLLIIFCNMRRPAHADIKTGYSIIMDGPVPLEEQCEYLSYDASQWEF 840
781 GTGVIAVFFWVLLLIIFCNMRRPAHADIKTGYSIIMDGPVPLEEQCEYLSYDASQWEF 840
841 PRERHLGRVLGYGAFGKVVESAFGTHKSSCDTVAVKMKEGATASEHRLMSKIL 900
841 PRERHLGRVLGYGAFGKVVESAFGTHKSSCDTVAVKMKEGATASEHRLMSKIL 900
901 THIGHNLNVNLLGACTKPGQPLMVIVFECKYGNLSNPLRAKRDAPSPCAEKSPQGRF 960
901 THIGHNLNVNLLGACTKPGQPLMVIVFECKYGNLSNPLRAKRDAPSPCAEKSPQGRF 960
961 RAMVELARLDRRRGSSDRVLFAFSTEGGARRASPDQAEIDLWLSPLTMEDLVCSYFQ 1020
961 RAMVELARLDRRRGSSDRVLFAFSTEGGARRASPDQAEIDLWLSPLTMEDLVCSYFQ 1020
1021 VARGMEFLASRCKIHRDLAARNILLSSEVDVYKICDFGLARDIYKDPDYVRKGSARLPKW 1080
1021 VARGMEFLASRCKIHRDLAARNILLSSEVDVYKICDFGLARDIYKDPDYVRKGSARLPKW 1080
1081 MAPESIFDKVYTTQSDVWSFGVLLWEIFSLGASPYPGVQINEEFCQRLRQDTRAPELA 1140
1081 MAPESIFDKVYTTQSDVWSFGVLLWEIFSLGASPYPGVQINEEFCQRLRQDTRAPELA 1140

QY 1021 VARGMEFLASRKC1HRDLAARNILLSESDVVKICDFGLARDIYKDPDYVRKGSARLPKW 1080
DB 1021 VARGMEFLASRKC1HRDLAARNILLSESDVVKICDFGLARDIYKDPDYVRKGSARLPKW 1080
QY 1081 MAPESIFDKVYTTOSDVMWFGVLLWEIFSLGASPPYGVQINEEFQORLDRDTRRAPELA 1140
DB 1081 MAPESIFDKVYTTOSDVMWFGVLLWEIFSLGASPPYGVQINEEFQORLDRDTRRAPELA 1140
QY 1141 TPAIRRMILNCWSDGPKARPAFSELVEILGDLLOGRGLQEEVEVCMAPRSSQSEEGSFS 1200
DB 1141 TPAIRRMILNCWSDGPKARPAFSELVEILGDLLOGRGLQEEVEVCMAPRSSQSEEGSFS 1200
QY 1201 QVSTMALHIAQDAEDSPSLQORHSLAARYNMVSPFCLARGAETRGSSRMRKTFEEFPM 1260
DB 1201 QVSTMALHIAQDAEDSPSLQORHSLAARYNMVSPFCLARGAETRGSSRMRKTFEEFPM 1260
QY 1261 TPTYKGSVDNQTDSGMVLASEEFEQIESRHRQESGFCCKGPGQNVAVTRAHPDSQGRRR 1320
DB 1261 TPTYKGSVDNQTDSGMVLASEEFEQIESRHRQESGFCCKGPGQNVAVTRAHPDSQGRRR 1320
QY 1321 RPERGARGQVYNSEYGEISELSEEDHCSPSARVTFEFTDINSY 1363
DB 1321 RPERGARGQVYNSEYGEISELSEEDHCSPSARVTFEFTDINSY 1363

RESULT 3
US-08-874-678-32
; Sequence 32, Application US/08874678
; Patent No. 5952199
; GENERAL INFORMATION:
; APPLICANT: Davis-Smyth, Terri L.
; APPLICANT: Chen, Helen H.
; APPLICANT: Presta, Leonard
; APPLICANT: Ferrara, Napoleone
; TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,678
; FILING DATE: HEREWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/643,839
; FILING DATE: 07-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-63291-1/WHID
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1363 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-874-678-32

Query Match 100.0%; Score 7211; DB 3; Length 1363;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1363; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MORGAAALCLRLMLCLGLDLGSLVSGYMTPTTLNITEESHVIDTGSLSISCRGHPLEWA 60
DB 1 MORGAAALCLRLMLCLGLDLGSLVSGYMTPTTLNITEESHVIDTGSLSISCRGHPLEWA 60
QY 61 WPGAQAPATGDKDSEDGTGVRCEGTDPAPYCKVLLLEHVNADTGSVCYKYIKARI 120
DB 61 WPGAQAPATGDKDSEDGTGVRCEGTDPAPYCKVLLLEHVNADTGSVCYKYIKARI 120
QY 121 EGTTAASSYVFVRDFEQPFINKPDTLLVNRKDAWMPCLVSPGLNVTLRSSSVLWPDG 180
DB 121 EGTTAASSYVFVRDFEQPFINKPDTLLVNRKDAWMPCLVSPGLNVTLRSSSVLWPDG 180
QY 181 QEVVMDRRGMLVSTPLLDHALYLOCTETGDDOFLSNPFLVHITGNELYDIQLPRKSL 240
DB 181 QEVVMDRRGMLVSTPLLDHALYLOCTETGDDOFLSNPFLVHITGNELYDIQLPRKSL 240
QY 241 ELLVGEKLVNCTWAEFNSGVTFDWDYPCQAERGKWPERRSQOHTLSILTIHNV 300
DB 241 ELLVGEKLVNCTWAEFNSGVTFDWDYPCQAERGKWPERRSQOHTLSILTIHNV 300
QY 301 SQHDLGSYVCKANNGIORPRESTEVIVHENPFIISVEWLGKPILEATAGDELKLPVKLAA 360
DB 301 SQHDLGSYVCKANNGIORPRESTEVIVHENPFIISVEWLGKPILEATAGDELKLPVKLAA 360
QY 361 YPPPEFQWYKDGKALSRHSPHALVLEKTEASTGYTTLALNLSAAGLRNLSLELVNVV 420
DB 361 YPPPEFQWYKDGKALSRHSPHALVLEKTEASTGYTTLALNLSAAGLRNLSLELVNVV 420
QY 421 PPOIHEKEASSPSYSRHSROALTCTAYGVPLPLSLQIOWHWRPMTCKMFAQRLRRQOQ 480
DB 421 PPOIHEKEASSPSYSRHSROALTCTAYGVPLPLSLQIOWHWRPMTCKMFAQRLRRQOQ 480
QY 481 DLMPQCDMRAVTTODAVNPTESLDTWTEFECKNTKSVKLVIQANVSAWYKCVVSNKV 540
DB 481 DLMPQCDMRAVTTODAVNPTESLDTWTEFECKNTKSVKLVIQANVSAWYKCVVSNKV 540
QY 541 QGDERLIYFYVTTIPDGFTTESKPESELLEGQPVLLSCQADSYKYEHLRWYRLNLSTLHD 600
DB 541 QGDERLIYFYVTTIPDGFTTESKPESELLEGQPVLLSCQADSYKYEHLRWYRLNLSTLHD 600
QY 601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLIPRVAPEHEGHVCEVODRRS 660
DB 601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLIPRVAPEHEGHVCEVODRRS 660
QY 661 HDKCHKKYLVSQALEAPRLTQNLTLVNVSDSLEMOCLVAGAHAPSIVWYKDERLLEE 720
DB 661 HDKCHKKYLVSQALEAPRLTQNLTLVNVSDSLEMOCLVAGAHAPSIVWYKDERLLEE 720
QY 721 KSGVDLADSNOKLSIQRVREDAGRYLCSVCNAKGCNVSSASVAVEGSDKGSMEIVILV 780
DB 721 KSGVDLADSNOKLSIQRVREDAGRYLCSVCNAKGCNVSSASVAVEGSDKGSMEIVILV 780
QY 781 GTGVIAVFFWVLLLLIFCNMRRPAHADIKTGLSIIMDPGEVPLEEQEVLSDYASQWMEF 840
DB 781 GTGVIAVFFWVLLLLIFCNMRRPAHADIKTGLSIIMDPGEVPLEEQEVLSDYASQWMEF 840
QY 841 PRERHLGRVLYGAFGVWASAFGIHKGSSCDTAVKMLKEGATASEHRLMSELKIL 900
DB 841 PRERHLGRVLYGAFGVWASAFGIHKGSSCDTAVKMLKEGATASEHRLMSELKIL 900
QY 901 THIGHNLNVNLLGACTKPOGLPQPLVIVFECKYGNLSNFRKRAKDAFSPCAEKSPQGRGF 960
DB 901 THIGHNLNVNLLGACTKPOGLPQPLVIVFECKYGNLSNFRKRAKDAFSPCAEKSPQGRGF 960
QY 961 RAWVELARLDRRRGSSDRVLFAFSPKTEGGARRASPDQAEADLWLSPLTMEDLVCSYFQ 1020
DB 961 RAWVELARLDRRRGSSDRVLFAFSPKTEGGARRASPDQAEADLWLSPLTMEDLVCSYFQ 1020

Query Match 99.3%; Score 7157; DB 2; Length 1363;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1354; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MORGALCLRLWCLGLDLGVSYSMTPTLNTITEESHVIDTGDLSLISCRQHPLEWA 60
DB 1 MORGALCLRLWCLGLDLGVSYSMTPTLNTITEESHVIDTGDLSLISCRQHPLEWA 60
QY 61 WPGAQAPATGDKDSEDGTGVRCEGTDPAPYCKVLLHEVHANDTGSVYCYKIKARI 120
DB 61 WPGAQAPATGDKDSEDGTGVRCEGTDPAPYCKVLLHEVHANDTGSVYCYKIKARI 120
QY 121 EGTAAASSYVFRDFEOPFINKPDTLLVNRKDAWVPCLVSIPLGNVTLRSQSSVLPDQ 180
DB 121 EGTAAASSYVFRDFEOPFINKPDTLLVNRKDAWVPCLVSIPLGNVTLRSQSSVLPDQ 180
QY 181 QEVVMDRRGMLVSTPLLDHALYLQCTTGWGDDQFLSNPFLVHITGNELYDIQLPRKSL 240
DB 181 QEVVMDRRGMLVSTPLLDHALYLQCTTGWGDDQFLSNPFLVHITGNELYDIQLPRKSL 240
QY 241 ELLVGEKLVNCTWAEFNSGVTFDWDYPCQAERGKWPERRSQOHTLSLTIHN 300
DB 241 ELLVGEKLVNCTWAEFNSGVTFDWDYPCQAERGKWPERRSQOHTLSLTIHN 300
QY 301 SHDLGSGVCKANNGIORFRESTEVIHNPFFISVENLKGPILEATAGDELVKPLVLA 360
DB 301 SHDLGSGVCKANNGIORFRESTEVIHNPFFISVENLKGPILEATAGDELVKPLVLA 360
QY 361 YPPEFOWYKDGKALSGRSHPHALVKEVTEASTGTTLWNSAAGLRNLSLELVNV 420
DB 361 YPPEFOWYKDGKALSGRSHPHALVKEVTEASTGTTLWNSAAGLRNLSLELVNV 420
QY 421 PPQIHEKEASSPSIYSRHSQALCTCTAYGVLPLSLIQHWRPMTCKNFADORSRRRQ 480
DB 421 PPQIHEKEASSPSIYSRHSQALCTCTAYGVLPLSLIQHWRPMTCKNFADORSRRRQ 480
QY 481 DLMPCQDRWRAVTTQDAVNPIESLDWTTFEVEGKNKTVSKLVIONANVSAMKCVSN 540
DB 481 DLMPCQDRWRAVTTQDAVNPIESLDWTTFEVEGKNKTVSKLVIONANVSAMKCVSN 540
QY 541 QGDERLIYFYVTTIPDGFTIESKSEELLEGOPVLLSCQADSKYKHEHLRWYRLNLSL 600
DB 541 QGDERLIYFYVTTIPDGFTIESKSEELLEGOPVLLSCQADSKYKHEHLRWYRLNLSL 600
QY 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLIPRAPEHEGHVYCEVQDRS 660
DB 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLIPRAPEHEGHVYCEVQDRS 660
QY 661 HDKCHKKYLSVQALEAPRLTONLTDLLVNVSDSLEMOCLVAGAHAPSIVWKDERLEE 720
DB 661 HDKCHKKYLSVQALEAPRLTONLTDLLVNVSDSLEMOCLVAGAHAPSIVWKDERLEE 720
QY 721 KSGVDLADSNKLSQRYREEDAGRYLCSVCNAKGCNVSSASVAVEGSEDKSGMEIVIL 780
DB 721 KSGVDLADSNKLSQRYREEDAGRYLCSVCNAKGCNVSSASVAVEGSEDKSGMEIVIL 780
QY 781 GTGVTAFFWVLLLIIFCNMRPRAHADIKTGYLSLIINDPGEVPLEEQEYLSYDASOWEF 840
DB 781 GTGVTAFFWVLLLIIFCNMRPRAHADIKTGYLSLIINDPGEVPLEEQEYLSYDASOWEF 840
QY 841 PRERHLGRVLYGAFGVVWASAFGIHKGSCDTPVAVKMLKEGATSEHRLMSLKIL 900
DB 841 PRERHLGRVLYGAFGVVWASAFGIHKGSCDTPVAVKMLKEGATSEHRLMSLKIL 900
QY 901 IHIGHNLNVNLLGACTKPOGLMVIYVEFCYKGNLSNFRALAKDAFSPCAEKSPQGRGF 960
DB 901 IHIGHNLNVNLLGACTKPOGLMVIYVEFCYKGNLSNFRALAKDAFSPCAEKSPQGRGF 960
QY 961 RAMVELARLDRRRPGSSDRVLFARFSPKTEGGARRASPOEAEDLWLSPLTMDLVCSYFQ 1020
DB 961 RAMVELARLDRRRPGSSDRVLFARFSPKTEGGARRASPOEAEDLWLSPLTMDLVCSYFQ 1020

RESULT 4

US-08-643-839-32
; Sequence 32, Application US/08643839
; Patent No. 6100071
; GENERAL INFORMATION:
; APPLICANT: Davis-Smyth, Terri L.
; APPLICANT: Chen, Helen H.
; APPLICANT: Presta, Leonard
; APPLICANT: Ferrara, Napoleone
; TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
; TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
; TITLE OF INVENTION: PRODUCTION
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/643,839
; FILING DATE: 07-MAY-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-63291/WHD
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELETYPE: 910 277299
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1363 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-643-839-32

Query Match

99.3%; Score 7157; DB 3; Length 1363;

Best Local Similarity 99.3%; Pred. No. 0;			
Matches 1354; Conservative 5; Mismatches 4; Indels 0; Gaps 0;			
QY	1	MORGAALCRLWLCIGLLDGLVSGYSMTPTLNITEESHVIDTGDSLSISCRGQHPLENA	60
DB	1	MORGAALCRLWLCIGLLDGLVSGYSMTPTLNITEESHVIDTGDSLSISCRGQHPLENA	60
QY	61	WPGAQAPATGDKDSEDGTGVVRCECTDARPYCKVLLHEVHANDTGSVCYKYIKARI	120
DB	61	WPGAQAPATGDKDSEDGTGVVRCECTDARPYCKVLLHEVHANDTGSVCYKYIKARI	120
QY	121	EGTTAASSYVVRDFEQPFINKPDTLLVNRKDAWVPCLVSTPGLNVTLRSSQSVLWPDG	180
DB	121	EGTTAASSYVVRDFEQPFINKPDTLLVNRKDAWVPCLVSTPGLNVTLRSSQSVLWPDG	180
QY	181	QEVVMDRRGMLYSTPLLDHALYLOQETTWGDDFLSNPFLVHITGNELYDIQLPRKSL	240
DB	181	QEVVMDRRGMLYSTPLLDHALYLOQETTWGDDFLSNPFLVHITGNELYDIQLPRKSL	240
QY	241	ELLVGEKLVNCTVWAEFNSGVTFDWDYFCKQAEKGVKWPERRSQOHTLSTLTHNV	300
DB	241	ELLVGEKLVNCTVWAEFNSGVTFDWDYFCKQAEKGVKWPERRSQOHTLSTLTHNV	300
QY	301	SOHDLGSYVCKANNRGRESTEVIHNPFFISVEWLGKPILEATAGDELVKLPVKLAA	360
DB	301	SOHDLGSYVCKANNRGRESTEVIHNPFFISVEWLGKPILEATAGDELVKLPVKLAA	360
QY	361	YPPEFOMYKDGKALSGRSHPHALVLKVEYTEASTGYTTLALWNSAAGLRNRSILELVNV	420
DB	361	YPPEFOMYKDGKALSGRSHPHALVLKVEYTEASTGYTTLALWNSAAGLRNRSILELVNV	420
QY	421	PPQIHEKEASSPSIYSHRSQALCTAYGVPLPLSTONHWRPWTCKMFAQRSRLRROQ	480
DB	421	PPQIHEKEASSPSIYSHRSQALCTAYGVPLPLSTONHWRPWTCKMFAQRSRLRROQ	480
QY	481	DLMPQCDRAWTTQDAVNPISLDTWTBEVECKNTVSKLVTONANVSAMYKCVVSNKV	540
DB	481	DLMPQCDRAWTTQDAVNPISLDTWTBEVECKNTVSKLVTONANVSAMYKCVVSNKV	540
QY	541	QODERLIYFVYTTIPDGFTIESKPSBELLEGQPVLLSCQADSYKYEHLRWYRLNLSTLHD	600
DB	541	QODERLIYFVYTTIPDGFTIESKPSBELLEGQPVLLSCQADSYKYEHLRWYRLNLSTLHD	600
QY	601	AHGNPLLLDCKNVHLFATPLAASLEEVAPCARHATLSLSPRVAPHEGHCYCEVODRRS	660
DB	601	AHGNPLLLDCKNVHLFATPLAASLEEVAPCARHATLSLSPRVAPHEGHCYCEVODRRS	660
QY	661	HDKHCHKKYLSVQALEAPRLTONLTDLLVNVSDSLEMOCLVACAHAPSIVWYKDERLLEE	720
DB	661	HDKHCHKKYLSVQALEAPRLTONLTDLLVNVSDSLEMOCLVACAHAPSIVWYKDERLLEE	720
QY	721	KSQVDLADSNOKLSIQVRREDAGRYLCSVCNNAKGCNVSSASVAVEGSDKGSNEIVILV	780
DB	721	KSQVDLADSNOKLSIQVRREDAGRYLCSVCNNAKGCNVSSASVAVEGSDKGSNEIVILV	780
QY	781	GTGVIAVFFVWVLLLLIFCNMRRPAHADIKTGYSLIIMDPCEVPLEEOCEYLSYDASQWEF	840
DB	781	GTGVIAVFFVWVLLLLIFCNMRRPAHADIKTGYSLIIMDPCEVPLEEOCEYLSYDASQWEF	840
QY	841	PRERHLGRVLGYGAFGKVVVYASAFGIHKGSSCDTVAVKMLKEGATASEHRLMSLKIL	900
DB	841	PRERHLGRVLGYGAFGKVVVYASAFGIHKGSSCDTVAVKMLKEGATASEHRLMSLKIL	900
QY	901	ITHGNHNVNLLGACTKPGQPLMVIVFCKYCNLSNLFRAKRDASFPCAESKSPGEORGRF	960
DB	901	ITHGNHNVNLLGACTKPGQPLMVIVFCKYCNLSNLFRAKRDASFPCAESKSPGEORGRF	960
QY	961	RAMVELARLDRRPGSSDRVLFAFARSKTEGGARRASPDQEAEDLWLSPLTMEDLVCSYFQ	1020
DB	961	RAMVELARLDRRPGSSDRVLFAFARSKTEGGARRASPDQEAEDLWLSPLTMEDLVCSYFQ	1020
QY	1021	VARGMEFLASRKCIIHRDLAARNILLSESDVVKICDFGLARDIYKDPDYVRKGSARLPLKW	1080
DB	1021	VARGMEFLASRKCIIHRDLAARNILLSESDVVKICDFGLARDIYKDPDYVRKGSARLPLKW	1080

RESULT 5

US-08-222-616-33
; Sequence 33, Application US/08222616
; Patent No. 5635177
; GENERAL INFORMATION:
; APPLICANT: Bennett, Brian D.
; APPLICANT: Goeddel, David
; APPLICANT: Lee, James M.
; APPLICANT: Matthews, William
; APPLICANT: Tsai, Siao Ping
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/222,616
; FILING DATE: 4-APR-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00586
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/826935
; FILING DATE: 22-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 821p2
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1298 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-222-616-33

1021 VARGMEFLASRKC IHRDLAARNILLSESDVVK ICDFGLARDIYKDPDYVRKGSARLPKW 1080
1021 VARGMEFLASRKC IHRDLAARNILLSESDVVK ICDFGLARDIYKDPDYVRKGSARLPKW 1080
1081 MAPESIFDKVYTTQSDVMSFGVLLWEIFSLGASPYPGVQINEEFCORLROGTRMRAPELA 1140
1081 MAPESIFDKVYTTQSDVMSFGVLLWEIFSLGASPYPGVQINEEFCORLROGTRMRAPELA 1140
1141 TPAIRRIHNCWSDPKRPAFSELVELTGLDGLAGRIQEEVEVCMAPRSQSSEGSFS 1200
1141 TPAIRRIHNCWSDPKRPAFSELVELTGLDGLAGRIQEEVEVCMAPRSQSSEGSFS 1200
1201 QVSTMALHIAQADAEDESPSLQRHSLAARYNNVSPFCCLARGAETRGSSRMKTFEEFPM 1260
1201 QVSTMALHIAQADAEDESPSLQRHSLAARYNNVSPFCCLARGAETRGSSRMKTFEEFPM 1260
1261 PPTTYKGSVDNOTDSGMVLASEEFEQIESRHRQESGF 1297
1261 PPTTYKGSVDNOTDSGMVLASEEFEQIESRHRQESGF 1297

RESULT 6

US-08-340-011-2
Sequence 2, Application US/08340011
Patent No. 5776755
GENERAL INFORMATION:
APPLICANT: Alitalo, et al.
TITLE OF INVENTION: FLT4, A NOVEL RECEPTOR TYROSIN KINASE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,011
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/959,951
FILING DATE: 09-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Thomas C.
REGISTRATION NUMBER: 36,989
REFERENCE/DOCKET NUMBER: 32267
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1298 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-011-2

Query Match 95.0%; Score 6852; DB 1; Length 1298;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.0%; Score 6852; DB 1; Length 1298;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MORGAAALCLRLWLCGLLDGLVSGYSMTPTNLNITEESHVIDTGDLSLISCRGQHPLEWA 60
1 MORGAAALCLRLWLCGLLDGLVSGYSMTPTNLNITEESHVIDTGDLSLISCRGQHPLEWA 60
61 WPGAQAPATGDKSDSDTGVVRCEGTDAKPYCKVLLHEVHANDTGSVVCYKIKARI 120
61 WPGAQAPATGDKSDSDTGVVRCEGTDAKPYCKVLLHEVHANDTGSVVCYKIKARI 120
121 EGTTAASVYVVRDFEOPFINKPDTLLVNRKDAWVPCVLSIPGLNVTLRSSSVLWPDG 180
121 EGTTAASVYVVRDFEOPFINKPDTLLVNRKDAWVPCVLSIPGLNVTLRSSSVLWPDG 180
181 QEVVMDRRGMLVSTPLLDHALYLQCTTWGDDQFLSNPFLVHITGNELYDIQLPRKSL 240
181 QEVVMDRRGMLVSTPLLDHALYLQCTTWGDDQFLSNPFLVHITGNELYDIQLPRKSL 240
241 ELLVGEKLVNCTVNAEFNSGVTFDWDYPCQAERGKWPERRSQOHTTSLTILHN 300
241 ELLVGEKLVNCTVNAEFNSGVTFDWDYPCQAERGKWPERRSQOHTTSLTILHN 300
301 SOHDLGSSVCKANNCIORFRESTEVI VHNPFISVENLKGPILEATAGDELKPLVKLAA 360
301 SOHDLGSSVCKANNCIORFRESTEVI VHNPFISVENLKGPILEATAGDELKPLVKLAA 360
361 YPPPEFQWKDGKALSGRHSHPALVKEVTEASTGTTLALWNSAAGLRNLSLELVV 420
361 YPPPEFQWKDGKALSGRHSHPALVKEVTEASTGTTLALWNSAAGLRNLSLELVV 420
421 PQIHEKASPSIYRHSRQALCTAGVPLPLSIQHWHPMTPCKMFQORSRLRRQOQ 480
421 PQIHEKASPSIYRHSRQALCTAGVPLPLSIQHWHPMTPCKMFQORSRLRRQOQ 480
481 DLMPCQROWRVTQDANVPITSLDTWTFEVEGKNKTVSKLVIONANVSAMKCVSNKV 540
481 DLMPCQROWRVTQDANVPITSLDTWTFEVEGKNKTVSKLVIONANVSAMKCVSNKV 540
541 GDERLIYFYVTIPDGTITKSESEELLEGOPVLLSCQADSYKYEHLRWYRLNLTLD 600
541 GDERLIYFYVTIPDGTITKSESEELLEGOPVLLSCQADSYKYEHLRWYRLNLTLD 600
601 AHGNPILLCKNVHLFATPLAASLEVAPGARHATLSLIPRAPEHEGHVYVCEVQDRS 660
601 AHGNPILLCKNVHLFATPLAASLEVAPGARHATLSLIPRAPEHEGHVYVCEVQDRS 660
661 HDKHKHKYLSVOALEAPRLTONLTDLLVNSDSLEMOCLVAGAHAPSIVMYKORLLEE 720
661 HDKHKHKYLSVOALEAPRLTONLTDLLVNSDSLEMOCLVAGAHAPSIVMYKORLLEE 720
721 KSGVDLADSNOKLSIORVREEDAGRYLCSVNAKGCNVSSASVAVEGSEDKGSMIEIV 780
721 KSGVDLADSNOKLSIORVREEDAGRYLCSVNAKGCNVSSASVAVEGSEDKGSMIEIV 780
781 GTGVIAVFWVWLLLLIFCNHRRPAHADIKTGYSLSIINDGPEVPLEEOCEYLSYDASQNEF 840
781 GTGVIAVFWVWLLLLIFCNHRRPAHADIKTGYSLSIINDGPEVPLEEOCEYLSYDASQNEF 840
841 PRELHLGRVLGVGAFCKVVEASAFQIHKGSSCDTVAVKMLKEGATASEHRLMSLKIT 900
841 PRELHLGRVLGVGAFCKVVEASAFQIHKGSSCDTVAVKMLKEGATASEHRLMSLKIT 900
901 IHGNHLNVNVLGACTKFGPLWTVFECKYGNLSNLFRAKRDAPSPCAEKSPQGRGF 960
901 IHGNHLNVNVLGACTKFGPLWTVFECKYGNLSNLFRAKRDAPSPCAEKSPQGRGF 960
961 RAMVELARLDRRPGSSDRVLFARFVKTEGGARRASPOQAEADLWLSPLTMDLVCSYFQ 1020
961 RAMVELARLDRRPGSSDRVLFARFVKTEGGARRASPOQAEADLWLSPLTMDLVCSYFQ 1020

QY 61 WPAQAEAPATGDKDSEDTGVVRCDEGTGDPYCKVLLHEVHANDTGSVYCYKIKARI 120
 DB 61 WPAQAEAPATGDKDSEDTGVVRCDEGTGDPYCKVLLHEVHANDTGSVYCYKIKARI 120
 QY 121 EGTAAASVYVVRDFEQPFINKPDTLLVNRKDAWVPCVLSIFGLNVTLSQSSVLPDQ 180
 DB 121 EGTAAASVYVVRDFEQPFINKPDTLLVNRKDAWVPCVLSIFGLNVTLSQSSVLPDQ 180
 QY 181 QEVVMDRRGMLVSTPLLDHALYQCTTGWGDDQFVLSNPFLVHTGNELYDIQLPRKSL 240
 DB 181 QEVVMDRRGMLVSTPLLDHALYQCTTGWGDDQFVLSNPFLVHTGNELYDIQLPRKSL 240
 QY 241 ELLVGEKLVNCTVWAEFNSGVTFDWDYPCQAERGKWPERRSQOHTTSLSSILTIHNV 300
 DB 241 ELLVGEKLVNCTVWAEFNSGVTFDWDYPCQAERGKWPERRSQOHTTSLSSILTIHNV 300
 QY 301 SHDLGSGVCKANNGIQRFRETEVIVHENPFTISVEWMLKGPILAEATAGDELVKLPVKLAA 360
 DB 301 SHDLGSGVCKANNGIQRFRETEVIVHENPFTISVEWMLKGPILAEATAGDELVKLPVKLAA 360
 QY 361 YPPPEFOMYKDGKALSGRHSPLALVKEVTEASTGTTLALWNSAAGLRNLSLELVNV 420
 DB 361 YPPPEFOMYKDGKALSGRHSPLALVKEVTEASTGTTLALWNSAAGLRNLSLELVNV 420
 QY 421 PPQIHEKEASSPSIYSRHSRQALCTAYGVPPLSLIQHWRPMTPKMFAORSRRRQOQ 480
 DB 421 PPQIHEKEASSPSIYSRHSRQALCTAYGVPPLSLIQHWRPMTPKMFAORSRRRQOQ 480
 QY 481 DLMPCRDWRVATVQDAVNPVIESLDTWTEFVEGKNKTVSKLVTONANVSAMYKCVSNKV 540
 DB 481 DLMPCRDWRVATVQDAVNPVIESLDTWTEFVEGKNKTVSKLVTONANVSAMYKCVSNKV 540
 QY 541 GODERLIYFYVTTPDGFTIESKSEELLEGQVLLSCQADSYKYEHLRWYRLNLSLTHD 600
 DB 541 GODERLIYFYVTTPDGFTIESKSEELLEGQVLLSCQADSYKYEHLRWYRLNLSLTHD 600
 QY 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLSTPRVAPEHEGHVYCEVQDRRS 660
 DB 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLSTPRVAPEHEGHVYCEVQDRRS 660
 QY 661 HDKCHKYKLSVQALEAPRLTQNTDLLVNSDSLEMOCLVAGAHAPSVIYVKDERLLEE 720
 DB 661 HDKCHKYKLSVQALEAPRLTQNTDLLVNSDSLEMOCLVAGAHAPSVIYVKDERLLEE 720
 QY 721 KSGVDLADSNOKLSIQRVREEDAGRYLCSVCNAKGVNSSASVAVESGDKGSMIEVTLV 780
 DB 721 KSGVDLADSNOKLSIQRVREEDAGRYLCSVCNAKGVNSSASVAVESGDKGSMIEVTLV 780
 QY 781 GTGVAVFFWVLLLIIFCNMRPAHADIKTGYLSIIMDPGEVPLEEOCEYLSDASQWEF 840
 DB 781 GTGVAVFFWVLLLIIFCNMRPAHADIKTGYLSIIMDPGEVPLEEOCEYLSDASQWEF 840
 QY 841 PRERHLGRVLGYAGFGVWEASAFGIHKGSSCDTAVVKMLKEGATASEHRALSELKIL 900
 DB 841 PRERHLGRVLGYAGFGVWEASAFGIHKGSSCDTAVVKMLKEGATASEHRALSELKIL 900
 QY 901 IHIGHNLNVNVLGACTKPGQLMVIVFEFCYKGNLSNPLAKRDAFSPCAEKSPQRGRF 960
 DB 901 IHIGHNLNVNVLGACTKPGQLMVIVFEFCYKGNLSNPLAKRDAFSPCAEKSPQRGRF 960
 QY 961 RAMVELARLDLRRRPGSSDRVLFARFSTKTEGGARRASPDQAEEDLWLSPLTHEDLVCSYFQ 1020
 DB 961 RAMVELARLDLRRRPGSSDRVLFARFSTKTEGGARRASPDQAEEDLWLSPLTHEDLVCSYFQ 1020
 QY 1021 VARGMEFLASRKCIHRDLAARNILLESDDVVKICDFGLARDIYKDPDYVRKGSARLPLKW 1080
 DB 1021 VARGMEFLASRKCIHRDLAARNILLESDDVVKICDFGLARDIYKDPDYVRKGSARLPLKW 1080
 QY 1081 MAPESIFDKVYTTQSDVNSFGVLLWEIFSLGASPYPGVQIINEEFCORLRDCTRMRAPELA 1140
 DB 1081 MAPESIFDKVYTTQSDVNSFGVLLWEIFSLGASPYPGVQIINEEFCORLRDCTRMRAPELA 1140

QY 1141 TPAIRRLMNCNSGDKPAPAFSELVEILGDLLOGRLQEEVEECVCMAPRSSQSEEGSFS 1200
 DB 1141 TPAIRRLMNCNSGDKPAPAFSELVEILGDLLOGRLQEEVEECVCMAPRSSQSEEGSFS 1200
 QY 1201 QVSIMALHIAQADAEDSPSLQRHSLAARYNNVSPGCLARGAETRGSSRMKTFEEFPM 1260
 DB 1201 QVSIMALHIAQADAEDSPSLQRHSLAARYNNVSPGCLARGAETRGSSRMKTFEEFPM 1260
 QY 1261 TPITYKGSVDNQTDGMLVASEEFEQIESRHRQESGF 1297
 DB 1261 TPITYKGSVDNQTDGMLVASEEFEQIESRHRQESGF 1297

RESULT 7

US-08-901-710-2
 : Sequence 2, Application US/08901710
 : Patent No. 6107046
 : GENERAL INFORMATION:
 : APPLICANT: Ailitalo, Kari
 : APPLICANT: Aprelikova, Olga
 : APPLICANT: Pajusola, Katri
 : APPLICANT: Armstrong, Elina
 : APPLICANT: Korhonen, Jaana
 : APPLICANT: Kaipainen, Arja
 : APPLICANT: Matikainen, Marja-Terttu
 : TITLE OF INVENTION: FL14, A RECEPTOR TYROSINE KINASE, AND USES
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 : STREET: 6300 Sears Tower, 233 South Wacker Drive
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: United States of America
 : ZIP: 60606-6402
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/901,710
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/340,011
 : FILING DATE: 14-NOV-1994
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/257,754
 : FILING DATE: 09-JUL-1994
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 07/959,951
 : FILING DATE: 09-OCT-1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Gass, David A.
 : REGISTRATION NUMBER: 38,153
 : REFERENCE/DOCKET NUMBER: 28113/33824
 : TELEPHONE: 312/474-6300
 : TELEFAX: 312/474-0448
 : TELEX: 25-3856
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1298 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-901-710-2

Query Match 95.0%; Score 6852; DB 3; Length 1298;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MORGAAALCLRLWLCLGLLDGLVSGYSMTPTLNITEESHVIDTGDLSLISCRGQHPLWA 60
DB 1 MORGAAALCLRLWLCLGLLDGLVSGYSMTPTLNITEESHVIDTGDLSLISCRGQHPLWA 60
QY 61 WPGAQAPATGDKDSEDGTVVROCEGTADARPKYKLLHEVHNDTGSYCYKYIKARI 120
DB 61 WPGAQAPATGDKDSEDGTVVROCEGTADARPKYKLLHEVHNDTGSYCYKYIKARI 120
QY 121 EGTAAASSYVVRDFEOPFINKPDTLLVNRKDAWVPCVLSIPCLNVTLSQSSVLPD 180
DB 121 EGTAAASSYVVRDFEOPFINKPDTLLVNRKDAWVPCVLSIPCLNVTLSQSSVLPD 180
QY 181 QEVVMDRRGMLVSTPLLDHALYLQCEETWGDQDFLNPFLVHITGNELYDIQLLPKSL 240
DB 181 QEVVMDRRGMLVSTPLLDHALYLQCEETWGDQDFLNPFLVHITGNELYDIQLLPKSL 240
QY 241 ELLVGEKLVNCTVWAEFNSGVTTDWDYPGKQAEGRGWPERRSQOHTHELSSILTIN 300
DB 241 ELLVGEKLVNCTVWAEFNSGVTTDWDYPGKQAEGRGWPERRSQOHTHELSSILTIN 300
QY 301 SOHDLGYSYCKANNGIORFESTEVIVHENPFI SVEMLKGPILFATAGDELVKLPVLA 360
DB 301 SOHDLGYSYCKANNGIORFESTEVIVHENPFI SVEMLKGPILFATAGDELVKLPVLA 360
QY 361 YPPPEFQYKDGKALSGRHSHPALVLEKEVTEASTGTTYTLWNLSAAGLRNLSLELV 420
DB 361 YPPPEFQYKDGKALSGRHSHPALVLEKEVTEASTGTTYTLWNLSAAGLRNLSLELV 420
QY 421 PPOIHEKASSPSYSRHSRQALCTAYGVPPLPSIOHHRPWTCKMFAORSRRRQO 480
DB 421 PPOIHEKASSPSYSRHSRQALCTAYGVPPLPSIOHHRPWTCKMFAORSRRRQO 480
QY 481 DLMPCRDWRVTTQDANVTIESLDTWTFEYEGNKNVSKLVIONANVSAMKCVSNKV 540
DB 481 DLMPCRDWRVTTQDANVTIESLDTWTFEYEGNKNVSKLVIONANVSAMKCVSNKV 540
QY 541 GODERLIYFYVTTIPDGTIESKPEELLEGOPVLLSCQADSYKYEHLRWYRLNLSLTH 600
DB 541 GODERLIYFYVTTIPDGTIESKPEELLEGOPVLLSCQADSYKYEHLRWYRLNLSLTH 600
QY 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLIPRVAPEHEGHYVCEVQDR 660
DB 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLIPRVAPEHEGHYVCEVQDR 660
QY 661 HDKCHKYKLVQALAPRLTQNTDLLVNSDSLEMQCLVAGAHAPSIVMYKDERLEE 720
DB 661 HDKCHKYKLVQALAPRLTQNTDLLVNSDSLEMQCLVAGAHAPSIVMYKDERLEE 720
QY 721 KSGVDLADSNOKLSIORVEDAGRYLCSVCNAKGCYNSSASVAVEGSKGMEIVTLV 780
DB 721 KSGVDLADSNOKLSIORVEDAGRYLCSVCNAKGCYNSSASVAVEGSKGMEIVTLV 780
QY 781 GTGVIAVFWFVLLLIIFCNRMPAHADIKTGYLSIIMDPGEVPLEEQEYLSYDASQW 840
DB 781 GTGVIAVFWFVLLLIIFCNRMPAHADIKTGYLSIIMDPGEVPLEEQEYLSYDASQW 840
QY 841 PRERHLGRVLYGAFGVKVEASAFIHKSGSCDTVAVKMLKEGATASEHRLMSKIL 900
DB 841 PRERHLGRVLYGAFGVKVEASAFIHKSGSCDTVAVKMLKEGATASEHRLMSKIL 900
QY 901 IHIGNHLNVNLLGACTKPGQPLMVIVFECKYGNLSNPLRAKRDASFCAEKSPEQRGR 960
DB 901 IHIGNHLNVNLLGACTKPGQPLMVIVFECKYGNLSNPLRAKRDASFCAEKSPEQRGR 960
QY 961 RAWVELARLDRRRPGSSDRVLFARFESKTEGGARRASPDQAEEDLWLSPLTHMEDLV 1020
DB 961 RAWVELARLDRRRPGSSDRVLFARFESKTEGGARRASPDQAEEDLWLSPLTHMEDLV 1020
QY 1021 VARGHEFLASRCKIHRDLAARNILLESDDVVKICDFGLARDIYKDPDVRKGSARLPKW 1080
DB 1021 VARGHEFLASRCKIHRDLAARNILLESDDVVKICDFGLARDIYKDPDVRKGSARLPKW 1080

QY 1081 MAPESIFDKVYTTQSDVMSEFVILLWEIFSLGASPYPCVQVINEEFCQRLDGTMRAPELA 1140
DB 1081 MAPESIFDKVYTTQSDVMSEFVILLWEIFSLGASPYPCVQVINEEFCQRLDGTMRAPELA 1140
QY 1141 TPAIRRLMNCWSDPKARPAFSELVELILGDLLOGLQEEVEECVMAPRSSQSESGSFS 1200
DB 1141 TPAIRRLMNCWSDPKARPAFSELVELILGDLLOGLQEEVEECVMAPRSSQSESGSFS 1200
QY 1201 QVSTMALHIAQADAEDSPSLQRHSLAARYYNNVSPGCLARGAETRGSSRMKTFEFP 1260
DB 1201 QVSTMALHIAQADAEDSPSLQRHSLAARYYNNVSPGCLARGAETRGSSRMKTFEFP 1260
QY 1261 TPITYKGSVDNQTDSGMVLASEEPEQIESRHRQESGF 1297
DB 1261 TPITYKGSVDNQTDSGMVLASEEPEQIESRHRQESGF 1297
RESULT 8
PCT-US95-04228-33
: Sequence 33, Application PC/TUS9504228
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: APPLICANT: Bennett, Brian D.
: APPLICANT: Goeddel, David
: APPLICANT: Lee, James M.
: APPLICANT: Matthews, William
: APPLICANT: Tsai, Siao Ping
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: palin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04228
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/222616
: FILING DATE: 04-APR-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Wendy M. Lee
: REGISTRATION NUMBER: 00,000
: REFERENCE/DOCKET NUMBER: 821P3PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 33:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1298 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
PCT-US95-04228-33

Query Match 95.0%; Score 6852; DB 4; Length 1298;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MORGAAALCLRLWLCLGLLDGLVSGYSMTPTLNITEESHVIDTGDLSLISCRGQHPLWA 60
DB 1 MORGAAALCLRLWLCLGLLDGLVSGYSMTPTLNITEESHVIDTGDLSLISCRGQHPLWA 60

Tue Jan 23 12:06:51 2001

```
QY 61 WPGAQAPATGDKDSEDGTGVVRDCEGTDPYCKVLLHEVHANDTGSVYCYKIKARI 120
Db 61 WPGAQAPATGDKDSEDGTGVVRDCEGTDPYCKVLLHEVHANDTGSVYCYKIKARI 120
QY 121 EGTAAASSVYFVDFEQPIFKPDTLLVNRKDAWVPCVLSIPGLNVLTRSGSSVLPDQ 180
Db 121 EGTAAASSVYFVDFEQPIFKPDTLLVNRKDAWVPCVLSIPGLNVLTRSGSSVLPDQ 180
QY 181 QEVVMDRRGMLVSTPLLDALYLQCEETWGDQDFLSNPFVHTGNELYDIQLPRKSL 240
Db 181 QEVVMDRRGMLVSTPLLDALYLQCEETWGDQDFLSNPFVHTGNELYDIQLPRKSL 240
QY 241 ELLVGEKLVNCTVWAEFNSGVTFDWDYFGKAERGKWPERRSOOTHTELSILTIHV 300
Db 241 ELLVGEKLVNCTVWAEFNSGVTFDWDYFGKAERGKWPERRSOOTHTELSILTIHV 300
QY 301 SQHDLGSGYVCKANNGIORFRESTEVIHNPFTISVEMLKGPILFATAGDELVKLPVKLAA 360
Db 301 SQHDLGSGYVCKANNGIORFRESTEVIHNPFTISVEMLKGPILFATAGDELVKLPVKLAA 360
QY 361 YPPEFQWYKDGKALSGRHSFHALVLKVEVTEASTGTITLALMNSAAGLRNLSLELVNV 420
Db 361 YPPEFQWYKDGKALSGRHSFHALVLKVEVTEASTGTITLALMNSAAGLRNLSLELVNV 420
QY 421 PPOIHEKEASSPSIYSHRSQALCTCTAYGVLPLSTQWHRPWTCKMFAQRLRRQOQ 480
Db 421 PPOIHEKEASSPSIYSHRSQALCTCTAYGVLPLSTQWHRPWTCKMFAQRLRRQOQ 480
QY 481 DLMPQCRDRAVTTQDANPIESLDVTWTEFEVGNKNTVSKLVTONANVSAMKYCVVSNKV 540
Db 481 DLMPQCRDRAVTTQDANPIESLDVTWTEFEVGNKNTVSKLVTONANVSAMKYCVVSNKV 540
QY 541 GQDERLIYFYTTIPDGFTTESKPEELLEGOPVLLSCQADSKYKHEHLRWYRLNLSLHD 600
Db 541 GQDERLIYFYTTIPDGFTTESKPEELLEGOPVLLSCQADSKYKHEHLRWYRLNLSLHD 600
QY 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLSPRVAPEHEGHVCEVQDRRS 660
Db 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLSPRVAPEHEGHVCEVQDRRS 660
QY 661 HDKCHKYLSVQALEAPRLTQNTDILLVNSDSLEMOCLVAGAHAPSIVVYKDERLLEE 720
Db 661 HDKCHKYLSVQALEAPRLTQNTDILLVNSDSLEMOCLVAGAHAPSIVVYKDERLLEE 720
QY 721 KSGVDLADSNOKLSIQRVREEDAGRYLCSVCNAKGCNVSSASVAVEGSDKGSMEIVILV 780
Db 721 KSGVDLADSNOKLSIQRVREEDAGRYLCSVCNAKGCNVSSASVAVEGSDKGSMEIVILV 780
QY 781 GTGVIAVFWVLLLIIFCMRRPAHADIKTGYLSIIMDPGEVPLEEQCEYLSYDASOWEF 840
Db 781 GTGVIAVFWVLLLIIFCMRRPAHADIKTGYLSIIMDPGEVPLEEQCEYLSYDASOWEF 840
QY 841 PRERHLGRVLYGAGFKVVEASAFGIHKGSSCDTVAVKMLKEGATASEHRALMSLKIL 900
Db 841 PRERHLGRVLYGAGFKVVEASAFGIHKGSSCDTVAVKMLKEGATASEHRALMSLKIL 900
QY 901 IHTGNHLNVNLLGACTKPGPLWIVPECKYGNLSNFLRAKRDFAFSCAEKSPQGRGF 960
Db 901 IHTGNHLNVNLLGACTKPGPLWIVPECKYGNLSNFLRAKRDFAFSCAEKSPQGRGF 960
QY 961 RAMVELARLDRRPGSSDRVLFAFESKTEGGARRASPDQEAEDLWLSPLTMEDLVCSYFQ 1020
Db 961 RAMVELARLDRRPGSSDRVLFAFESKTEGGARRASPDQEAEDLWLSPLTMEDLVCSYFQ 1020
QY 1021 VARGMEFLASRKC1HRDLAARNLILSESVDVKICDGLARDIYKQDPYVRKGSARLPLKW 1080
Db 1021 VARGMEFLASRKC1HRDLAARNLILSESVDVKICDGLARDIYKQDPYVRKGSARLPLKW 1080
QY 1081 MAPESIFDKVYTTQSDVNSFGVLLWEIFSLGASPYGVQVINEFCORLDRGTMRAPELA 1140
Db 1081 MAPESIFDKVYTTQSDVNSFGVLLWEIFSLGASPYGVQVINEFCORLDRGTMRAPELA 1140
QY 1141 TPAIRRLNLCWSDGPKARPAFSELVILGDLQGRGLQEEVEVCNAPRSSOSEEGSFS 1200
Db 1141 TPAIRRLNLCWSDGPKARPAFSELVILGDLQGRGLQEEVEVCNAPRSSOSEEGSFS 1200
QY 1201 QVSTMALHIAQADAEDSPPSLORHSLAARYNNVSPGCLARGAETRGSSRMKTFEFPFM 1260
Db 1201 QVSTMALHIAQADAEDSPPSLORHSLAARYNNVSPGCLARGAETRGSSRMKTFEFPFM 1260
QY 1261 TPTTYKGSVDNQTSGMVLASEEFEQIESRHRQESGF 1297
Db 1261 TPTTYKGSVDNQTSGMVLASEEFEQIESRHRQESGF 1297
RESULT 9
US-08-874-678-34
: Sequence 34, Application US/08874678
: Patent No. 5952199
: GENERAL INFORMATION:
: APPLICANT: Davis-Smyth, Terri L.
: APPLICANT: Chen, Helen H.
: APPLICANT: Presta, Leonard
: APPLICANT: Ferrara, Napoleone
: TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR
: TITLE OF INVENTION: ACTIVITY, THEIR USES AND PROCESSES FOR THEIR PRODUCTION
: NUMBER OF SEQUENCES: 48
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
: STREET: Four Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: United States
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/874.678
: FILING DATE: HEREWITH
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/643,839
: FILING DATE: 07-MAY-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Dreger, Walter H.
: REGISTRATION NUMBER: 24,190
: REFERENCE/DOCKET NUMBER: A-63291-1/MHD
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 34:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1368 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: protein
: MOLECULE TYPE: protein
: US-08-874-678-34
Query Match 94.3%; Score 6797.5; DB 2: Length 1368;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 1293; Conservative 23; Mismatches 46; Indels 7; Gaps 4;
QY 1 MORGALCLRLWLCGLLDGLVSGYSMTPTPLNI-TEESHVIDTGDLSLSISCRGHPLEW 59
Db 1 MORGALCLRLWLCGLLDGLVSGYSMTPTPLSIITTESHVIDTGDLSLSISCRGHPLEW 60
QY 60 AMPGAQAPATGDKDSEDGTGVVRDCEGTDPYCKVLLHEVHANDTGSVYCYKIKARI 119
Db 60 AMPGAQAPATGDKDSEDGTGVVRDCEGTDPYCKVLLHEVHANDTGSVYCYKIKARI 120
```


Db 121 IEGTTAAIYIFISDTRCPFEVMEYSEIPEI IHWTGRELIVPCRVTSNITVTLLKKFPLD 180
Qy 175 VLMPDQGVVMDORRGMLYSTPLLDHALYQCETTWGDQDFLSNPFVLHITGNELYDIOL 234
Db 181 TLIPDGKRIIWDKRKGFIIISNATYKIEIGLTLCEATVNGHLYKTN-YLTHRTGNELYDIOL 239
Qy 235 LPRKSLLELVGERKLVNCTVWAFNSGVTFDWDYPCQKQKRGKWPERRSQOHTHTLSL 294
Db 240 LPRKSLLELVGERKLVNCTVWAFNSGVTFDWDYPCQKQKRGKWPERRSQOHTHTLSL 299
Qy 295 LTHNVSQHDLSYVCKANNQORPRESTEVIVHENPFLSVLWLGKPILEATAGDELVLK 354
Db 300 LTHNVSQHDLSYVCKANNQORPRESTEVIVHENPFLSVLWLGKPILEATAGDELVLK 359
Qy 355 PVKLAAYPPPEFOWYKDGKALSGRHSFHALVLKVEVTEASTGYTTLALMNSAAGLRNLSL 414
Db 360 PVKLAAYPPPEFOWYKDGKALSGRHSFHALVLKVEVTEASTGYTTLALMNSAAGLRNLSL 419
Qy 415 ELVVNVPPQIHEKEASSPSYSRHSQALCTCTAYGVPLPLSLIOMHWRPMTCKMFAQRL 474
Db 420 ELVVNVPPQIHEKEASSPSYSRHSQALCTCTAYGVPLPLSLIOMHWRPMTCKMFAQRL 479
Qy 475 RRRQQDLMPQCRDRAVTTQDAVNPIESLDWTTEVEGKNKTVSKLVIONANVSAMYK 534
Db 480 RRRQQDLMPQCRDRAVTTQDAVNPIESLDWTTEVEGKNKTVSKLVIONANVSAMYK 539
Qy 535 VSNKVGQDERLIYFYVTTPDGTIESKPSLEELLEGQPVLLSCQADSYKYEHLRWYRLN 594
Db 540 VSNKVGQDERLIYFYVTTPDGTIESKPSLEELLEGQPVLLSCQADSYKYEHLRWYRLN 599
Qy 595 LSTLHDAGNPLLLDCKNVHLFATPIAASLEEVAPGARHATLSLSIPRAPEHEGHVCE 654
Db 600 LSTLHDAGNPLLLDCKNVHLFATPIAASLEEVAPGARHATLSLSIPRAPEHEGHVCE 659
Qy 655 VQDRSHDKHCHKYLSVQALEAPRLTONLTDLNVNVSQLEMOCLVAGAHAFSIYVYKD 714
Db 660 VQDRSHDKHCHKYLSVQALEAPRLTONLTDLNVNVSQLEMOCLVAGAHAFSIYVYKD 719
Qy 715 ERLLSEKSGVDLADSNOKLSIQRVREDAAGYLCVNCNAGCVNNSASVAVEGSEDKGSM 774
Db 720 ERLLSEKSGVDLADSNOKLSIQRVREDAAGYLCVNCNAGCVNNSASVAVEGSEDKGSM 779
Qy 775 EIVILVGTGVIAVFFWVLLLIIFCNMRRPAHADIKTGYLSIIMDPGEVPLEEOCEYLSYD 834
Db 780 EIVILVGTGVIAVFFWVLLLIIFCNMRRPAHADIKTGYLSIIMDPGEVPLEEOCEYLSYD 839
Qy 835 ASOWEFPREHLHGRVLGYGAFGKVVESAFG IHKGSSECDTVAVKMLKEGATASEHRLM 894
Db 840 ASOWEFPREHLHGRVLGYGAFGKVVESAFG IHKGSSECDTVAVKMLKEGATASEHRLM 899
Qy 895 SELKILIHGNHLNVNLLGACTKPGQPLMVIVFECKYGNLSNPLRAKRAOAFSPCAEKSP 954
Db 900 SELKILIHGNHLNVNLLGACTKPGQPLMVIVFECKYGNLSNPLRAKRAOAFSPCAEKSP 959
Qy 955 EORGRFRAMVELALDRRRPSSDRVLFAFBSKTEGGARRASPDQEAEDLWLSPLTHMEDL 1014
Db 960 EORGRFRAMVELALDRRRPSSDRVLFAFBSKTEGGARRASPDQEAEDLWLSPLTHMEDL 1019
Qy 1015 VCYSFQVARGMEFLASRCKIHRDLAARNILLSSESDVYKICDFGLARDIYKDPDYVRKGS 1074
Db 1020 VCYSFQVARGMEFLASRCKIHRDLAARNILLSSESDVYKICDFGLARDIYKDPDYVRKGS 1079
Qy 1075 RLPLKWAAPESIFDKVYTTQSDVWSFGVLLWEIFSLGASYPGVQVINEEFQORLDRGTRM 1134
Db 1080 RLPLKWAAPESIFDKVYTTQSDVWSFGVLLWEIFSLGASYPGVQVINEEFQORLDRGTRM 1139
Qy 1135 RAPELATPAIRIMLNCWSGDPKARPAFSELVEILGDLGRLQEEVEECVMAPRSSQSS 1194
Db 1140 RAPELATPAIRIMLNCWSGDPKARPAFSELVEILGDLGRLQEEVEECVMAPRSSQSS 1199
Qy 1195 EBGSPQSVMTHALHTAQADAEDSPSLQRIHSLAARYYNWVSFPGCLARGAETRGSSRHKT 1254
Db 1200 EBGSPQSVMTHALHTAQADAEDSPSLQRIHSLAARYYNWVSFPGCLARGAETRGSSRHKT 1259

Qy 1255 FEEFPMPTTYKSGVDNQDTSQMWLASBEEFOIESHRHROESGFSCKGPGQNVAVTRAHPD 1314
Db 1260 FEEFPMPTTYKSGVDNQDTSQMWLASBEEFOIESHRHROESGFSCKGPGQNVAVTRAHPD 1319
Qy 1315 SOGRRRRRGARGGOVFNSEYSGELSPSEEDHCSPSARVTFFTDNSY 1363
Db 1320 SOGRRRRRGARGGOVFNSEYSGELSPSEEDHCSPSARVTFFTDNSY 1368

RESULT 11

US-08-874-678-33

: Sequence 33, Application US/08874678

: Patent No. 5952199

: GENERAL INFORMATION:

: APPLICANT: Davis-Smyth, Terri L.

: APPLICANT: Chen, Helen H.

: APPLICANT: Presta, Leonard

: APPLICANT: Ferrara, Napoleone

: TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR

: TITLE OF INVENTION: ACTIVITY, THEIR USES AND PROCESSES FOR THEIR PRODUCTION

: NUMBER OF SEQUENCES: 48

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

: STREET: Four Embarcadero Center, Suite 3400

: CITY: San Francisco

: STATE: California

: COUNTRY: United States

: ZIP: 94111-4187

: COMPUTER READABLE FORM:

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent In Release #1.0, Version #1.30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/874,678

: FILING DATE: HEREWITH

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 08/643,839

: FILING DATE: 07-MAY-1996

: ATTORNEY/AGENT INFORMATION:

: NAME: Dreger, Walter H.

: REGISTRATION NUMBER: 24,190

: REFERENCE/DOCKET NUMBER: A-63291-1/WHD

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (415) 781-1989

: TELEFAX: (415) 398-3249

: TELEX: 910 277299

: INFORMATION FOR SEQ ID NO: 33:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 1362 amino acids

: TYPE: amino acid

: STRANDEDNESS: unknown

: TOPOLOGY: unknown

: MOLECULE TYPE: protein

US-08-874-678-33

Query Match 83.3%

Best Local Similarity 84.1%

Matches 1153; Conservative 61; Mismatches 140; Indels 17; Gaps 6;

Qy 1 MQGGAALCLRLWLCLGLLDGLVSGYSMTPTLNTITEESHVIDTGDLSLSISCRQHLEWA 60
Db 1 MQGGAALCLRLWLCLGLLDGLVSGYSMTPTLNTITEESHVIDTGDLSLSISCRQHLEWA 60
Qy 61 WPGAQEAAPATGDKDSEDTGVVRCEGTDAHPYCKVLLHLLHEVHANDTGSYVCYKIKARI 120
Db 61 -----LPMVWSKESERLSITKSACGRNGKQFCSTLTNTAQANHGTGYSKYLAVPTSK 114
Qy 121 EGTTAASSYVVRDPEOPFINK-----PDTLLVNRKDAWVPCLVSIPLGNVTLRS-OSSV 175
Db 121 EGTTAASSYVVRDPEOPFINK-----PDTLLVNRKDAWVPCLVSIPLGNVTLRS-OSSV 175

Db 115 KKETESAIYIFISDTGRPFVEMYSEIPEIHHMTEGRELVIPCRVTSNPTVTLKKFPDPT 174
Qy 176 LWPQGVVWDDRRGMLVSTPLLDALYLQCTETMGDDQDFLNSPLFVHTGNELVDIQLL 235
Db 175 LIPDKRIIWDSSRKGFIIISNATYKEIGLITCEATVNGHLYKTN-YLTHROTWTIIDVQIS 233
Qy 236 PRKSLELLVGEKILVNLCTVMAEFNSGVTFDWDYPGQAEGRKWVPERNSSQOOTHLS --- 292
Db 234 TPRPVKLLRGHTLVNLCTATTPLNTRVQMTWSYDPEKKNRAS--VRRRIQDSNHSANIEV 291
Qy 293 SILTIHNSOHDLSGVCKANNGIQRFRESTEVIYHNFISVEMWLGKPILEATAGDELY 352
Db 292 SVLTIDKMNQKRGVLCVRSRPSFKSVSRVHNFISVEMWLGKPILEATAGDELY 351
Qy 353 KLPVKLAAYPPPEFOWYKDGKALSGRHSFHALVKEVTEASTGTITLALWNSAAGLRNRI 412
Db 352 KLPVKLAAYPPPEFOWYKDGKALSGRHSFHALVKEVTEASTGTITLALWNSAAGLRNRI 411
Qy 413 SLELVVWPQIHEKEASSPSIYSRHSRQALTCTAYGVPLPLSIOHWRPWTPTCKMFAOR 472
Db 412 SLELVVWPQIHEKEASSPSIYSRHSRQALTCTAYGVPLPLSIOHWRPWTPTCKMFAOR 471
Qy 473 SLRRQODLWPCQCRWRVTTQDANPIESLDTWTEFVEGKNKTSKLVIONANSAMY 532
Db 472 SLRRQODLWPCQCRWRVTTQDANPIESLDTWTEFVEGKNKTSKLVIONANSAMY 531
Qy 533 KCVVSNKVGQDERLIYFVTTIPDGFTIESKPSSEELLEGOPVLLSCQADSKYKHEHLRWYR 592
Db 532 KCVVSNKVGQDERLIYFVTTIPDGFTIESKPSSEELLEGOPVLLSCQADSKYKHEHLRWYR 591
Qy 593 LNLSTLHDACNPLLLDCKNHLFATPLAASLEEVAPGARHATLSIPRAVEHEGHYV 652
Db 592 LNLSTLHDACNPLLLDCKNHLFATPLAASLEEVAPGARHATLSIPRAVEHEGHYV 651
Qy 653 CEVQDRSHDKHCHKKLSVQALEAPRLTQNLTLNVLNVSDESLEMOCLVAGAHAPSIVWY 712
Db 652 CEVQDRSHDKHCHKKLSVQALEAPRLTQNLTLNVLNVSDESLEMOCLVAGAHAPSIVWY 711
Qy 713 KDERLLEKSGVDLADSNQKLSQVRREEDAGRYLCSVCNAKGCNVSSASVAVEGSEDKG 772
Db 712 KDERLLEKSGVDLADSNQKLSQVRREEDAGRYLCSVCNAKGCNVSSASVAVEGSEDKG 771
Qy 773 SMEIVLVGTGVIAVFFWVLLLIIFCNMRPRAHADIKTGYSLIIMDPGEVPLEEQCEYLS 832
Db 772 SMEIVLVGTGVIAVFFWVLLLIIFCNMRPRAHADIKTGYSLIIMDPGEVPLEEQCEYLS 831
Qy 833 YDASOWEPFRERHLGRVLYGAFKGVVEASAFGIHKGSSCDTVAVKMLKEGATASEHRA 892
Db 832 YDASOWEPFRERHLGRVLYGAFKGVVEASAFGIHKGSSCDTVAVKMLKEGATASEHRA 891
Qy 893 LMSELKILIHGNILNVLNLLGACTKPGPLMWIVFECKYGNLSNLRKARDAFSPCAEK 952
Db 892 LMSELKILIHGNILNVLNLLGACTKPGPLMWIVFECKYGNLSNLRKARDAFSPCAEK 951
Qy 953 SPEQGRFRANVELARDRRPGRSSDRVLFARFSKTEGGARRASPDQEAEDLWLSPLTME 1012
Db 952 SPEQGRFRANVELARDRRPGRSSDRVLFARFSKTEGGARRASPDQEAEDLWLSPLTME 1011
Qy 1013 DLVCYSFQVARGMEFLASRCKIHRDLAARNILLESVDVKICDFGLARDIYKOPDYVRKG 1072
Db 1012 DLVCYSFQVARGMEFLASRCKIHRDLAARNILLESVDVKICDFGLARDIYKOPDYVRKG 1071
Qy 1073 SARLPLKMAPESTFDKVTYTQSDVWSFGVLLWEIFSLGASPYPGVQINEEFQORLDRGT 1132
Db 1072 SARLPLKMAPESTFDKVTYTQSDVWSFGVLLWEIFSLGASPYPGVQINEEFQORLDRGT 1131
Qy 1133 RRAPELATPAIRIMLNCWSDGPKARPAFSELVEILGLDLOGRGLQEEVEVCMAPRSSQ 1192
Db 1132 RRAPELATPAIRIMLNCWSDGPKARPAFSELVEILGLDLOGRGLQEEVEVCMAPRSSQ 1191
Qy 1193 SSEEGSFQVSTMALHTAQADAEDSPSLQRHSLAARYYNNWSPFGCLARGAETRGSSRM 1252
Db 1192 SSEEGSFQVSTMALHTAQADAEDSPSLQRHSLAARYYNNWSPFGCLARGAETRGSSRM 1251

Qy 1253 KTFEEFPMPTPTYKGSVDNQTDSGMVLASEEPEQIESRHRQESGFSCKGPGQNVAVTRAH 1312
Db 1252 KTFEEFPMPTPTYKGSVDNQTDSGMVLASEEPEQIESRHRQESGFSCKGPGQNVAVTRAH 1311
Qy 1313 PDSQGRRRRPERGARGQGVFNSEYGEISEEDHCSPSARVTFFTONS 1363
Db 1312 PDSQGRRRRPERGARGQGVFNSEYGEISEEDHCSPSARVTFFTONS 1362

RESULT 12
US-08-643-839-33
Sequence 33, Application US/08643839
Patent No. 6100071
GENERAL INFORMATION:
APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Presta, Leonard
APPLICANT: Ferrara, Napoleone
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
TITLE OF INVENTION: PRODUCTION
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/643,839
FILING DATE: 07-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-63291/WH0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 1362 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-643-839-33

Query Match 83.3%; Score 6005.5; DB 3; Length 1362;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 1153; Conservative 61; Mismatches 140; Indels 17; Gaps 6;

Qy 1 MORGAAALRLWLCLGLDLGVSGYSMTPTLNTIESHVITDGLSLISCRQHPLEWA 60
Db 1 MORGAAALRLWLCLGLDLGVSGYSMTPTLSLKTQHTIMOAGOTLHLOCRGEAAHKS 60
Qy 61 WPGAQAPATGDKDSEDTGVVRDCECTDARPYCKVLLHLEVHANTDGSVCYKYKARI 120
Db 61 -----LPEMVSKESERLSITKSACGRNCKQFCSTLTNTAQANHTGFTSCKYLAVPISK 114
Qy 121 EGTIAASSYVVFDEQPFINK----PDTLLVNRKDMVPCVLSIPGLNVLRLS--QSSV 175
Db 115 KKETESAIYIFISDTGRPFVEMYSEIPEIHHMTEGRELVIPCRVTSNPTVTLKKFPDPT 174

```

DB 1252 KTFEEFPTPTTYKSGVDNQTDSGMVLASEEFOIESRHRQESGFCCKPGQGNVATRAH 1311
QY 1313 PDSQGRRRRPERGARGGOVFYNSEYGELSPSEEDHCSPSARVYFTTDSNY 1363
DB 1312 PDSQGRRRRPERGARGGOVFYNSEYGELSPSEEDHCSPSARVYFTTDSNY 1362

RESULT 13
US-08-874-678-3
; Sequence 3, Application US/08874678
; Patent No. 5952199
; GENERAL INFORMATION:
; APPLICANT: Davis-Smyth, Terri L.
; APPLICANT: Chen, Helen H.
; APPLICANT: Presta, Leonard
; APPLICANT: Ferrara, Napoleone
; TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR
; TITLE OF INVENTION: ACTIVITY, THEIR USES AND PROCESSES FOR THEIR PRODUCTION
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,678
; FILING DATE: HERESWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/643,839
; FILING DATE: 07-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Dregger, Walter H.
; REGISTRATION NUMBER: 24,190
; REFERENCE/DOCKET NUMBER: A-63291-1/WHO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 777 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-874-678-3

Query Match 57.3%; Score 4135; DB 2; Length 777;
Best Local Similarity 99.9%; Pred. NO. 0;
Matches 775; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MORGAAALCLRLWLCLGLLDGLVSGYSMTPTPLNITEESHVIDTGDLSLISCRGQHPLEWA 60
DB 1 MORGAAALCLRLWLCLGLLDGLVSGYSMTPTPLNITEESHVIDTGDLSLISCRGQHPLEWA 60
QY 61 WPGAQEA PATGDKDSED TG VVRDC EGT DARP YCKVLLH EVHANDTGS YVCYKYIKARI 120
DB 61 WPGAQEA PATGDKDSED TG VVRDC EGT DARP YCKVLLH EVHANDTGS YVCYKYIKARI 120
QY 121 ECTTAASSYVFVRFOEPFINKPOTLLVNRKDAWVPCLVSIPLAVNTLRSSSVLWPDG 180
DB 121 ECTTAASSYVFVRFOEPFINKPOTLLVNRKDAWVPCLVSIPLAVNTLRSSSVLWPDG 180

```

181 QEVWDDRRGMLVSTPLLDALYLQCTTWGDDFLSNPFLVHITGNELYDIQLPRKSL 240
181 QEVWDDRRGMLVSTPLLDALYLQCTTWGDDFLSNPFLVHITGNELYDIQLPRKSL 240
241 ELLVGEKLVNCTVNAEFGNSGVTDFWDYPGKQAEKGVKWPERRSOOTHTLSSILTIHNV 300
241 ELLVGEKLVNCTVNAEFGNSGVTDFWDYPGKQAEKGVKWPERRSOOTHTLSSILTIHNV 300
301 SQHDLGSGYVCKANNGIQRFESTEVIVHENPFISVEMWLGKPILEATAGDELKVLPLKLA 360
301 SQHDLGSGYVCKANNGIQRFESTEVIVHENPFISVEMWLGKPILEATAGDELKVLPLKLA 360
361 YPPPEFQWYKDKGKALSGRHSRPHALVKEVTEASTGTITLALNWSAAGLRNLSLELVNV 420
361 YPPPEFQWYKDKGKALSGRHSRPHALVKEVTEASTGTITLALNWSAAGLRNLSLELVNV 420
421 PQIHEKEASSPSIYSRHSRQALCTAYGVPLPSIQHWRPWTCKMFAQRSLRRQOO 480
421 PQIHEKEASSPSIYSRHSRQALCTAYGVPLPSIQHWRPWTCKMFAQRSLRRQOO 480
481 DLMQCRDMRAVTTQDAVNPISLDTWTEFVEGKNTVSKLVIONANVSAMKCVVSNKV 540
481 DLMQCRDMRAVTTQDAVNPISLDTWTEFVEGKNTVSKLVIONANVSAMKCVVSNKV 540
541 GQDERLIYFVVTTPDGFTIESKPSEELLEGOPVLLSCQADSYKYEHLRWYRLNLSLTHD 600
541 GQDERLIYFVVTTPDGFTIESKPSEELLEGOPVLLSCQADSYKYEHLRWYRLNLSLTHD 600
601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLSIPRVAPEHEGHYVCEVODRRS 660
601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLSIPRVAPEHEGHYVCEVODRRS 660
661 HDKCHHKYLSVOALEAPRLTQNLTLNVSDSLSEMOCLVAGAHAPSIVWYKDERLLEE 720
661 HDKCHHKYLSVOALEAPRLTQNLTLNVSDSLSEMOCLVAGAHAPSIVWYKDERLLEE 720
721 KSGVDLADSNQKLSIQRVREEDAGRYLCSVCNAKGCNVSSASVAVEGSEDKGSMEV 776
721 KSGVDLADSNQKLSIQRVREEDAGRYLCSVCNAKGCNVSSASVAVEGSEDKGSMEV 776

RESULT 14
US-08-643-839-3
Sequence 3, Application US/08643839
Patent No. 6100071
GENERAL INFORMATION:
APPLICANT: Davis-Smyth, Terri L.
APPLICANT: Chen, Helen H.
APPLICANT: Presta, Leonard
APPLICANT: Ferrara, Napoleone
TITLE OF INVENTION: NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
TITLE OF INVENTION: GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
TITLE OF INVENTION: PRODUCTION
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Rohbach, Test, Albrighton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/643,839
FILING DATE: 07-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190

REFERENCE/DOCKET NUMBER: A-63291/WHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 777 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-643-839-3

Query Match 57.3% Score 4135 DB 3 Length 777
Best Local Similarity 99.9% Pred. No. 0
Matches 775 Conservative 1 Mismatches 0 Indels 0 Gaps 0
QY 1 MORGALCLRLWLCGLLDGLVSGYSMTPTLNIITEESHVIDTGDLSISCRGQHPLEWA 60
DB 1 MORGALCLRLWLCGLLDGLVSGYSMTPTLNIITEESHVIDTGDLSISCRGQHPLEWA 60
QY 61 WPGAQEPATGDKDSEDTCVVRDCEGTDPYCKVLLHEVHANDTGSVCYKYIKARI 120
DB 61 WPGAQEPATGDKDSEDTCVVRDCEGTDPYCKVLLHEVHANDTGSVCYKYIKARI 120
QY 121 EGTTAASSYVVFVDFEQPFINKPDTLLVNRKDMWVPCVLSIPGLNVTLRSSQSVLWPDG 180
DB 121 EGTTAASSYVVFVDFEQPFINKPDTLLVNRKDMWVPCVLSIPGLNVTLRSSQSVLWPDG 180
QY 181 QEVWDDRRGMLVSTPLLDALYLQCTTWGDDFLSNPFLVHITGNELYDIQLPRKSL 240
DB 181 QEVWDDRRGMLVSTPLLDALYLQCTTWGDDFLSNPFLVHITGNELYDIQLPRKSL 240
QY 241 ELLVGEKLVNCTVNAEFGNSGVTDFWDYPGKQAEKGVKWPERRSOOTHTLSSILTIHNV 300
DB 241 ELLVGEKLVNCTVNAEFGNSGVTDFWDYPGKQAEKGVKWPERRSOOTHTLSSILTIHNV 300
QY 301 SQHDLGSGYVCKANNGIQRFESTEVIVHENPFISVEMWLGKPILEATAGDELKVLPLKLA 360
DB 301 SQHDLGSGYVCKANNGIQRFESTEVIVHENPFISVEMWLGKPILEATAGDELKVLPLKLA 360
QY 361 YPPPEFQWYKDKGKALSGRHSRPHALVKEVTEASTGTITLALNWSAAGLRNLSLELVNV 420
DB 361 YPPPEFQWYKDKGKALSGRHSRPHALVKEVTEASTGTITLALNWSAAGLRNLSLELVNV 420
QY 421 PQIHEKEASSPSIYSRHSRQALCTAYGVPLPSIQHWRPWTCKMFAQRSLRRQOO 480
DB 421 PQIHEKEASSPSIYSRHSRQALCTAYGVPLPSIQHWRPWTCKMFAQRSLRRQOO 480
QY 481 DLMQCRDMRAVTTQDAVNPISLDTWTEFVEGKNTVSKLVIONANVSAMKCVVSNKV 540
DB 481 DLMQCRDMRAVTTQDAVNPISLDTWTEFVEGKNTVSKLVIONANVSAMKCVVSNKV 540
QY 541 GQDERLIYFVVTTPDGFTIESKPSEELLEGOPVLLSCQADSYKYEHLRWYRLNLSLTHD 600
DB 541 GQDERLIYFVVTTPDGFTIESKPSEELLEGOPVLLSCQADSYKYEHLRWYRLNLSLTHD 600
QY 601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLSIPRVAPEHEGHYVCEVODRRS 660
DB 601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLSIPRVAPEHEGHYVCEVODRRS 660
QY 661 HDKCHHKYLSVOALEAPRLTQNLTLNVSDSLSEMOCLVAGAHAPSIVWYKDERLLEE 720
DB 661 HDKCHHKYLSVOALEAPRLTQNLTLNVSDSLSEMOCLVAGAHAPSIVWYKDERLLEE 720
QY 721 KSGVDLADSNQKLSIQRVREEDAGRYLCSVCNAKGCNVSSASVAVEGSEDKGSMEI 776
DB 721 KSGVDLADSNQKLSIQRVREEDAGRYLCSVCNAKGCNVSSASVAVEGSEDKGSMEV 776
RESULT 15

US-08-810-116-8
Sequence 8, Application US/08810116
Patent No. 5766860
GENERAL INFORMATION:
APPLICANT: Terman, Bruce I.
TITLE OF INVENTION: Identification of a No. 5766860e Human Growth
Factor Receptor
TITLE OF INVENTION: Factor Receptor
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/810,116
FILING DATE: 25-FEB-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/930,548
FILING DATE: 23-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31,298-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1356 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-810-116-8

Query Match 38.2%; Score 2753; DB 1: Length 1356;
Best Local Similarity 43.9%; Pred. No. 5e-220;
Matches 606; Conservative 205; Mismatches 480; Indels 88; Gaps 26;

QY 1 MORGALCLRLWLCGLLDGLVS--GYSMPTPTLITEESHVIDTGDLSLSCRGQHPLE 58
Db 1 MESKYLVAVALWLCVETRAASVGLPSVLDLPLRSLQKIDILITKANTLQICRGQRDL 60
QY 59 WAPGQAEPATGDKDSEDGTGWRDC-EGTDARPYCKVLLLHEVHANDTGSVYCYK 117
Db 61 WLPNNQSG-----SEQRVEVTECSGL----FCKTLTPKVIQNDTGAYKCFYR--- 106
QY 118 ARIEGTTAASSVVFVRDEQFINKPDT-----LLVNRKDAWVCLVSTIPGLNVTL-- 169
Db 107 ---ETDLASVIVVQVDRSPFIASVSDQHGVIYITENKNTVVPICLGSINLSVLC 163
QY 170 RSQSSVLPDGEVVDVDRRGLVSTPLLDHALYLQCETTWGDQDFLSNPELVHITCNEL 229
Db 164 RYPERKFPDGNRI\$WDSKKGFTIPSYMISYAGWVCEAKINDESQSIYIVVWVGRI 223
QY 230 YDIQLPRKSLLELGEKLVNCTWAEFNSGVTFDWDYPCQAERGKWPERRSQOHTI 289
Db 224 YDVVLSPSHGIELSVGEKLVNCTARTNLNVIDFNPSPSKHKKLVNRDLKTOGSG 283
QY 290 ELS---SILTIIHNSQHDLSVYCKANNGIQRFRESTEVIVHENPFISVEMKLPILEAT 346
Db 284 EMKFLSTLITDGTFRDQGLYTCASSGLMTKKNSTFVRVHEKPFVAFGSMESLVEAT 343
QY 347 AGDELVKLPVLAAYPPPEFQWKGKALSGRH---SPHALVLEKVEASTGTYYTLALNW 403

Db 344 VG-ERVRIPAKYLGYPPEIKWYKNGIPLESNHTIKACHVLTIMVSESDTGNVYILTN 402
QY 404 SAAGLRNRISLELVVWVPPQIHEKEASSP-SYVSRHSROALTCTAYGVPLPLSIQHWMP 462
Db 403 PISKEKOSHWVSVLVVYVPPQIGEKSLISPVDSYQYGTQTTLCTVYVIPPHHIHWML 462
QY 463 WIPCKNFAQSRSLRRQOQDLMPOCRDMRAVTTQDAVNPIESLDTWTFEVEGNKTVSKLV 522
Db 463 EEE---ANEPQAVSVTNYP-CEEMRSVDEFGGKIEVKNKNOFALIEGKNKTVSTLV 518
QY 523 IQANVSAMKYCVVSNKVGQDERLIYFYVTTIPDGTIESKPSSEELLGQPVLLSCADS 582
Db 519 IQAANVSALYKCEAVNKVGRGERVISHVTRGPE---ITLQPMQPTQESVSLWCTADR 575
QY 583 KYEHLRWYRLNLSTLDAHGPNLLLDCKNV---HLFATPLAASLEEVAPGARHATLSL 638
Db 576 STFENLTWYKLGPOPLPHVGEPLTPVCKNLDLTWKLNMATFMSNSTNDI-----LIM 627
QY 639 SIPRVAPEHEGHVYCEVQDRSHDKCHKKYLSVQALEAPRLTONLTDLLVNVSDSLEMQ 698
Db 628 ELKNASLQDQGDYVCLAQDRKTKRHCVRQLTVLERVAPTITGNLENQTSIGESTEV 687
QY 699 CLVAGAHAPSTWYKDERLLEKSGVDLADSNOKLSIQRVREEDAGRYLCSVCNAKGVN 758
Db 688 CTASGNPPQIMHFKDNETLVEDSGIVLKGCRNLTIRRVKKEDEGLYTCQACSVLCAK 747
QY 759 SSASVAVESGDKSMEIVLVGTGVIAVFWLILLIFCNMRRPAHADIKTGYLSLIMD 818
Db 748 VEAFFIEGAQKTNLEIILVGTVTIAMEFWLLVLTILGTVKRANGELKGTGYLSLIMD 807
QY 819 PCEVPLEOCEYLSVDASOWEPEPRERLHLGVLGYGAFGVKVEASAFGIHKGSSCDTVAV 878
Db 808 PDELPLDEHCERLPYDASKWEFPDRNLNGLKPLGRGAFQGEIADAFIDKATCTAVAV 867
QY 879 KMLKEGATSEHRAHLMSELKILIHIGHNLNVNLLGACTKPGGLWVIVECKYGNLSNP 938
Db 868 KMLKEGATSEHRAHLMSELKILIHIGHNLNVNLLGACTKPGGLWVIVECKYGNLSNP 927
QY 939 LRARDAFSPCAEKSPEORGRFRA---NVELARLDRRRPGSSDRVLFARFARFARFAR 995
Db 928 LRSKNEFVYPYKTKG----ARFQKQYVGAIPYDLKR--RLDSTISSASSASSGFVEEK 981
QY 996 S----PDQEA-EDLWLSPLTMEDLVYCSFQVARGHEFLASRKCIRHDLAARNILLSVDV 1050
Db 982 SLSDVEEEAEPEDLYKDFLTLEHLICYSFQVAKGHEFLASRKCIRHDLAARNILLSSEKV 1041
QY 1051 VKICDFGLARDIYKDPDYVRKGSARLPLKMWAPESIFDKVYTTQSDVNSFGVLLWEIFSL 1110
Db 1042 VKICDFGLARDIYKDPDYVRKGDARLPLKMWAPETIFDRVYTIQSDVNSFGVLLWEIFSL 1101
QY 1111 GASPYGVOINEEFCORLDRGTRMRAPELATPAIRIRMLNCWNGDPKARPAFSELVEILG 1170
Db 1102 GASPYGVKIDEEFCRLKEGTRMRAPDVTPPEMVTQMLDCWHGEPESORPTFSELVEHIG 1161
QY 1171 DLLOGRGLQEEVEVCMAPRS-SOSSEEGSFQSVTMAHIAQADAEDSPSLQRHSLAAR 1229
Db 1162 NLLQANAQDQGDYIVLPISLTSMEEDSGLSLPTSPVSCHEEEVEVCDP-----KPH 1213
QY 1230 YNNVWSPFCCLARGAETGRSSRMKTPEEFPM-TPTTYKSGVDNQTDSGMVLASEEPEQIE 1288
Db 1214 YDMTAGISQYLQNSKRSRPSVKTFEDIPLEPEVKVIPPDDNQTDSGMVLASEELKLE 1273
QY 1289 SRHQESGFCCKPGONAVTRAHPDSQRRRRPERGARGGVFYNSEYSELSESEED 1347
Db 1274 DRTKLSFSGMVPK-----SRESVASEGSNQ-----TSYGOSGYHSDTDTDTTVVSE 1323

Search completed: January 23, 2001, 05:03:22
Job time: 445 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 23, 2001, 05:01:52 ; Search time 65.64 Seconds
(without alignments)
1409.942 Million cell updates/sec

Title: US-09-375-248-2
Perfect score: 7211
Sequence: 1 MORGAAALCLRLWLCGLLDG.....SEEDHCSPSARVTFITDINSY 1363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6852	95.0	1298	2 A48999	protein-tyrosine k
2	6320	87.6	1363	2 I58375	protein-tyrosine k
3	4919.5	68.2	1379	2 JC4954	vascular endotheli
4	2804	38.9	1348	2 JC4953	vascular endotheli
5	2804	38.9	1348	2 S51656	protein-tyrosine k
6	2753	38.2	1356	2 JC1402	protein-tyrosine k
7	2669	37.0	1367	2 A41228	receptor tyrosine
8	2469	34.2	1333	2 I78875	protein-tyrosine k
9	2463	34.2	1338	2 S09982	Fit-1 tyrosine kin
10	2444.5	33.9	1336	2 I60598	embryonic receptor
11	2437.5	33.8	1330	2 S49010	endothelial kinase
12	1283.5	17.8	367	2 B56598	platelet-derived g
13	1163.5	16.1	1089	1 PFHUGA	platelet-derived g
14	1151	16.0	1088	1 PFRTGA	platelet-derived g
15	1125	15.6	1089	1 S33727	platelet-derived g
16	1122.5	15.6	1087	2 I51552	platelet-derived g
17	1084	15.0	1106	1 PFHUGB	platelet-derived g
18	1052.5	14.6	1098	1 PFMSRB	macrophage colony
19	1051	14.6	975	2 T30816	protein-tyrosine k
20	1044	14.5	978	1 A49814	protein-tyrosine k
21	1042.5	14.5	976	1 TVHUKT	protein-tyrosine k
22	1033.5	14.3	975	1 TVMSKT	protein-tyrosine k
23	1024	14.2	977	2 I45877	protein-tyrosine k
24	1021.5	14.2	960	2 J45877	protein-tyrosine k
25	1018	14.1	954	2 I51703	C-kit-related kina
26	1012	14.0	1048	2 T30815	platelet-derived g
27	967	13.4	980	1 TVCTMD	macrophage colony-
28	956	13.3	941	1 TVMVMD	protein-tyrosine k
29	954.5	13.2	972	1 TVHUMD	macrophage colony-

ALIGNMENTS

RESULT 1

A48999

protein-tyrosine kinase (EC 2.7.1.112) flt4 precursor - human

N:Alternate names: class III receptor tyrosine kinase FLT4; fms-like tyrosine kinase

C:Species: Homo sapiens (man)

C>Date: 21-Jan-1994 #sequence.revision 18-Nov-1994 #text_change 20-Jun-2000

C:Accession: A48999; A44930; G02316; S36130; A42010

R:Pajusola, K.; Aprelikova, O.; Korhonen, J.; Kaipainen, A.; Pertovaara, L.; Alitalo,

Cancer Res. 52, 5738-5743, 1992

A:Title: FLT4 receptor tyrosine kinase contains seven immunoglobulin-like loops and 1

A:Reference number: A48999; MUID:93007958

A:Accession: A48999

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1298 <PAJ>

A:CROSS-references: PIDN: AAB23636; PID: g257352

A:Experimental source: HEL erythroleukemia cells

A:Note: sequence extracted from NCBI backbone (NCBIP:115335)

R:Pajusola, O.; Pajusola, K.; Partanen, J.; Armstrong, E.; Alitalo, R.; Bailey, S.K.

Cancer Res. 52, 746-748, 1992

A:Title: FLT4, a novel class III receptor tyrosine kinase in chromosome 5q33-qter.

A:Reference number: A44930; MUID:92119639

A:Accession: A44930

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 761-1190 <APB>

A:CROSS-references: GB:X58203; NID:g31433

A:Note: sequence extracted from NCBI backbone (NCBIP:78155)

R:Wood, W.I.

submitted to the EMBL Data Library, December 1995

A:Reference number: H01039

A:Accession: G02316

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1298 <WOO>

A:CROSS-references: EMBL:U43143; NID:g1150990; PIDN:AAA85215.1; PID:g1150991

R:Galland, F.

submitted to the EMBL Data Library, December 1992

A:Reference number: S36130

A:Accession: S36130

A:Molecule type: mRNA

A:Residues: 1-23, 'D', '25-744, 'P', '746-751, 'RP', '754-889, 'Q', '891-1127, 'V', '1129-1145, 'H', '1

A:CROSS-references: EMBL:X69878; NID:g297049; PIDN:CAA49505.1; PID:g297050

R:Galland, F.; Karamysheva, A.; Mattei, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.

Genomics 13, 475-478, 1992

A:Title: Chromosomal localization of FLT4, a novel receptor-type tyrosine kinase gene

A:Reference number: A42010; MUID:92307693

A:Accession: A42010

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 776-889, 'Q', '891-1127, 'V', '1129-1145, 'H', '1147-1163, 'D', '1165-1200 <GAL2>

C:Genetics:

A:Gene: GDB:FLT4
 A:Cross-references: GDB:128732; OMIM:136352
 A:Map position: 5q34-5q35
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
 C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; transmembrane
 F:1-23/Domain: signal sequence status predicted <SIG>
 F:24-1298/Product: protein-tyrosine kinase FLT4 status predicted <NAT>
 F:843-1176/Domain: protein kinase homology <KIN>
 F:851-859/Region: protein kinase ATP-binding motif

Query Match 95.0%; Score 6852; DB 2; Length 1298;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQRGAALCLRLWLCGLLDGLVSGYSMTPTLNITEESHVIDTGDLSLSCRGQPLEWA 60
 DB 1 MQRGAALCLRLWLCGLLDGLVSGYSMTPTLNITEESHVIDTGDLSLSCRGQPLEWA 60
 QY 61 WPGAQAPATGDKDSEDTGVVRDCEGTDPYCKVLLHEVHNDTGSVYCYIKARI 120
 DB 61 WPGAQAPATGDKDSEDTGVVRDCEGTDPYCKVLLHEVHNDTGSVYCYIKARI 120
 QY 121 EGTAASSVYVFRDFEOPFFINKPDTLLVNRKDAWVPCVLSIPGLNVLRSQSVLPDG 180
 DB 121 EGTAASSVYVFRDFEOPFFINKPDTLLVNRKDAWVPCVLSIPGLNVLRSQSVLPDG 180
 QY 181 QEVVMDRRGMLVSTPLLDHALYLCQETTWGDDFLSNPFLVHITGNELYDIQLPRKSL 240
 DB 181 QEVVMDRRGMLVSTPLLDHALYLCQETTWGDDFLSNPFLVHITGNELYDIQLPRKSL 240
 QY 241 ELLVGEKLVNCTVWAEFNSGVTFDMDYPGKQAEGRKWVPERSSQOHTFELSILTHV 300
 DB 241 ELLVGEKLVNCTVWAEFNSGVTFDMDYPGKQAEGRKWVPERSSQOHTFELSILTHV 300
 QY 301 SOHDLGSYVCKANNIGORFESTEVIVHNPFTSVEMLKGPILEATAGDELKPLVKLAA 360
 DB 301 SOHDLGSYVCKANNIGORFESTEVIVHNPFTSVEMLKGPILEATAGDELKPLVKLAA 360
 QY 361 YPPPEFOWYKDGKALSGRHSFHALVUKVEATSTGYITLALWNSAAGLRNLSLELVVNV 420
 DB 361 YPPPEFOWYKDGKALSGRHSFHALVUKVEATSTGYITLALWNSAAGLRNLSLELVVNV 420
 QY 421 PPQTHEKEASPSYSHRSQALCTCTAYGVPLPLSIQWHRWPTCKMPAQRSLRRQOQ 480
 DB 421 PPQTHEKEASPSYSHRSQALCTCTAYGVPLPLSIQWHRWPTCKMPAQRSLRRQOQ 480
 QY 481 DLMPQCRDRAVTTQDAVNPISLDITWTEFVEGKNKTYSKLVIONANVSAMKYCVVSNKV 540
 DB 481 DLMPQCRDRAVTTQDAVNPISLDITWTEFVEGKNKTYSKLVIONANVSAMKYCVVSNKV 540
 QY 541 GQDERLIYFYTTIPDGFTEKSEELLEGQPVLLSCQADSKYKHEHLRYRLNLSLTHD 600
 DB 541 GQDERLIYFYTTIPDGFTEKSEELLEGQPVLLSCQADSKYKHEHLRYRLNLSLTHD 600
 QY 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLSPRVAPHEGHVCEVODRRS 660
 DB 601 AHGNPLLDCKNVHLFATPLAASLEEVAPGARHATLSLSPRVAPHEGHVCEVODRRS 660
 QY 661 HDKCHKYLSVQALEAPRLTQNTDLLVNVSDSLEMOCLVAGAHAPSIVWYKDERLLEE 720
 DB 661 HDKCHKYLSVQALEAPRLTQNTDLLVNVSDSLEMOCLVAGAHAPSIVWYKDERLLEE 720
 QY 721 KSGVDLADSNKLSIQVRREDAGRYLCSVCNAKGCNVSSASVAVESGDKGMEIVILY 780
 DB 721 KSGVDLADSNKLSIQVRREDAGRYLCSVCNAKGCNVSSASVAVESGDKGMEIVILY 780
 QY 781 GTGVIAVEFFWLLLIIFCNMRPRAHDIKTGYLSIINDPGEVPLEOCEYLSYDASQWEF 840
 DB 781 GTGVIAVEFFWLLLIIFCNMRPRAHDIKTGYLSIINDPGEVPLEOCEYLSYDASQWEF 840
 QY 841 PRERHLGRVLGYGAFGKVVVEASAFGIHGSSCDTAVKMLKEGATASEHRLMSELKIL 900

DB 841 PRERHLGRVLGYGAFGKVVVEASAFGIHGSSCDTAVKMLKEGATASEHRLMSELKIL 900
 QY 901 IHIGNHLNVNLLGACITKPOGPLMVIIEFCKYGNLSNPLRAKRDAFSPCAEKSPQGRF 960
 DB 901 IHIGNHLNVNLLGACITKPOGPLMVIIEFCKYGNLSNPLRAKRDAFSPCAEKSPQGRF 960
 QY 961 RAMVELARLDRRRPGSSDRVLFAFESKTEGARRASPDQAEADLWLSPLTMEDLVCTSPQ 1020
 DB 961 RAMVELARLDRRRPGSSDRVLFAFESKTEGARRASPDQAEADLWLSPLTMEDLVCTSPQ 1020
 QY 1021 VARGMEFLASRKCITHRDLAARNILLSESVDVVKICDFGLARDIYKDPYVRKGSARLPLK 1080
 DB 1021 VARGMEFLASRKCITHRDLAARNILLSESVDVVKICDFGLARDIYKDPYVRKGSARLPLK 1080
 QY 1081 MAPESIFDKVYTTQSDVMSFGVLLWEIFSLGASPYGVQINEEFQRLRGTNRAPELA 1140
 DB 1081 MAPESIFDKVYTTQSDVMSFGVLLWEIFSLGASPYGVQINEEFQRLRGTNRAPELA 1140
 QY 1141 TPAIRRMILNCWSDGPKARPAFSELVEITLGLQGRGLQEEVEVCMAPRSSQSEDSFS 1200
 DB 1141 TPAIRRMILNCWSDGPKARPAFSELVEITLGLQGRGLQEEVEVCMAPRSSQSEDSFS 1200
 QY 1201 QVSTNALHIAQDAEDSPSLQRLSHLAARYNMYVSPGCLARGAETRGSSSRMKTFFEEFP 1260
 DB 1201 QVSTNALHIAQDAEDSPSLQRLSHLAARYNMYVSPGCLARGAETRGSSSRMKTFFEEFP 1260
 QY 1261 PPTTYKGSVDNQTDGSMVLASEEFEQIESRHRQESGF 1297
 DB 1261 PPTTYKGSVDNQTDGSMVLASEEFEQIESRHRQESGF 1297
 RESULT 2
 158375
 protein-tyrosine kinase (EC 2.7.1.112) flt4 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 04-Feb-2000
 C:Accession: 158375; B42010
 R:Galland, F.; Karamysheva, A.; Mattel, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.
 R:Finerty, H.; Kelleher, K.; Morris, G.E.; Bean, K.; Merberg, D.M.; Kriz, R.; Morris
 Oncogene 8, 2293-2298, 1993
 A:Title: Molecular cloning of murine FLT and FLT4.
 A:Reference number: 158375; MUID:93330572
 A:Accession: 158375
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1363 <RES>
 A:Cross-references: GB:L07296; NID:9293780; PID:AAA40077.1; PID:9293781
 R:Galland, F.; Karamysheva, A.; Mattel, M.G.; Rosnet, O.; Marchetto, S.; Birnbaum, D.
 Genomics 13, 475-478, 1992
 A:Title: Chromosomal localization of FLT4, a novel receptor-type tyrosine kinase gene
 A:Reference number: A42010; MUID:92307693
 A:Accession: B42010
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: DNA
 A:Residues: 1033-1072 <GAL>
 C:Genetics:
 A:Gene: FLT4
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
 F:843-1176/Domain: protein kinase homology <KIN>
 F:851-859/Region: protein kinase ATP-binding motif
 Query Match 87.6%; Score 6320; DB 2; Length 1363;
 Best Local Similarity 87.5%; Pred. No. 2.4e-282;
 Matches 1192; Conservative 68; Mismatches 103; Indels 0; Gaps 0;
 QY 1 MQRGAALCLRLWLCGLLDGLVSGYSMTPTLNITEESHVIDTGDLSLSCRGQPLEWA 60
 DB 1 MQRGAALCLRLWLCGLLDGLVSGYSMTPTLNITEESHVIDTGDLSLSCRGQPLEWA 60
 QY 61 WPGAQAPATGDKDSEDTGVVRDCEGTDPYCKVLLHEVHNDTGSVYCYIKARI 120
 DB 61 WPGAQAPATGDKDSEDTGVVRDCEGTDPYCKVLLHEVHNDTGSVYCYIKARI 120

QY 1201 QVSTMALHIAQDAEDSPSLOHSLAARYNNVSPFGLARGAETRGSRMRKTFEEFPM 1260
 DB 1201 QASTTALHTEADADSPSMECHSLAARYNCVSPFGLARGTKTPGSRMRKTFEELPM 1260
 QY 1261 TPITYKSDNODTSDGMLASEEFOIESHRHROESGCKPGQONVAVTRAHPDSDGRRR 1320
 DB 1261 TPITYKSDNODTSDGMLASEEFOIESHRHROESGCKPGQOHMDIPRGHPDPOGRRR 1320
 QY 1321 RPERGARGOVFNSEYSELSEEDHCSPSARVTFITDINSY 1363
 DB 1321 RPTQAGGKRVFNSEYSEVSPCTEGDCPCPSAGSTFFADSSY 1363

RESULT 3
 JC4954
 vascular endothelial growth factor receptor 2 precursor - Japanese quail
 N:Alternate names: Quail endothelial kinase 2; Quek 2
 C:Species: Coturnix coturnix japonica (Japanese quail)
 C:Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 24-Sep-1999
 C:Accession: JC4954
 R:Eichmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.
 Gene 174, 3-8, 1996
 A:Title: Molecular cloning of Quek 1 and 2, two quail vascular endothelial growth fac
 A:Reference number: JC4953; MUID:97017121
 A:Accession: JC4954
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-1379 <IC>
 A:Cross-references: EMBL:X83287; NID:q619865; PIDN:CAA58267.1; PID:e203815; PID:g1707
 C:Comment: This protein is an endothelial-specific receptor and binds vascular endoth
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
 C:Keywords: ATP; embryo; growth factor receptor; transmembrane protein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:789-810/Domain: transmembrane #status predicted <TM>
 F:856-1188/Domain: protein kinase homology <KIN>
 F:864-872/Region: protein kinase ATP-binding motif

Query Match 68.2%; Score 4919.5; DB 2; Length 1379;
 Best Local Similarity 68.2%; Pred. No. 3.5e-218;
 Matches 940; Conservative 158; Mismatches 246; Indels 35; Gaps 13;

QY 9 LRLNLCLGLD--GLVSGYSMTPTLTNTEESHVDTGDSLSISCRGQHPLEWAPGAQE 66
 DB 7 LPLNLGLVSEADLVSSISMTPTLTNTEESHVDTGDSLSISCRGQHPLEWAPGAQE 66
 QY 67 APATGDK-----DSEDGTG---VRDCEGT DARPYCKVLLLEHVEANDTGSTVCTYKIK 117
 DB 67 TPVEGRRNRNSQPQRPVPGAGNPEDECGTKPKYCKVLVLTESQANDTGYTHCYKID 126
 QY 118 ARIEGTTAASVYVDFPEOPFINKPDTLLVNRKDMVPCVLPGLNVLRSOSSVLM 177
 DB 127 AKIEGTAVSAVIFVDFPEOPFINKPDTLLVNRKDMVPCVLPGLNVLRSOSSVLM 186
 QY 178 PDGEVMDRRGMVSTPLLDHALYLOQETTWGQDFLSNPFVHTITGNELYDLOLPR 237
 DB 187 PDRTTFMDNKGQVPTQLIRDSLVQCTVIDNKFKNFIIHAGIAGLYDLYPK 246
 QY 238 KSELLVGLKLVNCTVWAEFNSGVTFDWDYDYGKQAEKFWPERSQOHTHELSILT 297
 DB 247 KAMELVGKLVNCTVWAEFNSGVTFDWDYDYGKQAEKFWPERSQOHTHELSILT 306
 QY 298 HNVSDHLDGVCYVCKANNGIQRFESTEVIVHENPFIISVEMKGPILAEATAGDELVKLPVK 357
 DB 307 HNVSDHLDGVCYVCKANNGIQRFESTEVIVHENPFIISVEMKGPILAEATAGDELVKLPVK 366
 QY 358 LAAYPPPEFQWYKDGKALSGRSPHALVLEKVEYTAEGTGTTLALNWSAGLRNLSLELV 417
 DB 367 VVAYPQDFQWYKAGK-LIPKQSSQSNQKDVAAEHAGITVTLVLRNLVGLKRLISQLI 425
 QY 418 VNVPPQIHEKASSPSIYSRHSRQALTCYAGVPLPLSIQWHRPMTCKHFAQRLSLR-- 475
 DB 426 VNVPPRIHEKASSPSIYSRHSRQALTCYAGVPLPLSIQWHRPMTCKHFAQRLSLR-- 485

QY 121 EGTAAASVYVDFPEOPFINKPDTLLVNRKDMVPCVLPGLNVLRSOSSVLM 180
 DB 121 EGTAAASVYVDFPEOPFINKPDTLLVNRKDMVPCVLPGLNVLRSOSSVLM 180
 QY 181 QVWDDRRGLVSTPLLDHALYLOQETTWGQDFLSNPFVHTITGNELYDLOLPR 240
 DB 181 QVWDDRRGLVSTPLLDHALYLOQETTWGQDFLSNPFVHTITGNELYDLOLPR 240
 QY 241 ELLVGEKLVNCTVWAEFNSGVTFDWDYDYGKQAEKFWPERSQOHTHELSILT 300
 DB 241 ELLVGEKLVNCTVWAEFNSGVTFDWDYDYGKQAEKFWPERSQOHTHELSILT 300
 QY 301 SOHDLGVCYVCKANNGIQRFESTEVIVHENPFIISVEMKGPILAEATAGDELVKLPVK 360
 DB 301 SOHDLGVCYVCKANNGIQRFESTEVIVHENPFIISVEMKGPILAEATAGDELVKLPVK 360
 QY 361 YPPPEFQWYKDGKALSGRSPHALVLEKVEYTAEGTGTTLALNWSAGLRNLSLELV 420
 DB 361 YPPPEFQWYKDGKALSGRSPHALVLEKVEYTAEGTGTTLALNWSAGLRNLSLELV 420
 QY 421 PPOIHEKASSPSIYSRHSRQALTCYAGVPLPLSIQWHRPMTCKHFAQRLSLR 480
 DB 421 PPOIHEKASSPSIYSRHSRQALTCYAGVPLPLSIQWHRPMTCKHFAQRLSLR 480
 QY 481 DLMPQCRDMRAVTTQDAVNPVIESLDTWTFEVEGKNTVSKLVQANVSAKYCVSNKV 540
 DB 481 DLMPQCRDMRAVTTQDAVNPVIESLDTWTFEVEGKNTVSKLVQANVSAKYCVSNKV 540
 QY 541 GQERLIYFYVTTIPDGFSTESPELLEGGQVPLTSCQADSKYKHLRWYRLNSTLHD 600
 DB 541 GQERLIYFYVTTIPDGFSTESPELLEGGQVPLTSCQADSKYKHLRWYRLNSTLHD 600
 QY 601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLIPVAPHEGHYCEVQDRRS 660
 DB 601 AHGNPLLLDCKNVHLFATPLAASLEEVAPGARHATLSLIPVAPHEGHYCEVQDRRS 660
 QY 661 HDRHCKKLYSVQALEAPRLTQNTDLLVNSDSLEMOCLVAGAHAPSIVWYKDERLLEE 720
 DB 661 HDRHCKKLYSVQALEAPRLTQNTDLLVNSDSLEMOCLVAGAHAPSIVWYKDERLLEE 720
 QY 721 KSGVDLADSNQKLSIQVRREEDAGRYLCSVCNAKGCYVNSASVAVEGSEDKGMEIVLV 780
 DB 721 KSGVDLADSNQKLSIQVRREEDAGRYLCSVCNAKGCYVNSASVAVEGSEDKGMEIVLV 780
 QY 781 GTGVIAVFFVWLLLIIFCNMRPAHADIKTGYSIIMDPGEVPLEEOCEYLSYDASOWEF 840
 DB 781 GTGVIAVFFVWLLLIIFCNMRPAHADIKTGYSIIMDPGEVPLEEOCEYLSYDASOWEF 840
 QY 841 PRERHLGRVLGAGFGRVVEASAFHGKSSCDTVAVKMLKEGATASEHRALMSELKIL 900
 DB 841 PRERHLGRVLGAGFGRVVEASAFHGKSSCDTVAVKMLKEGATASEHRALMSELKIL 900
 QY 901 IHIGHNLVNVNLLGACTPQGLVAVYVEFCYKGNLSNFRKARDASPCAEKSPQGRGF 960
 DB 901 IHIGHNLVNVNLLGACTPQGLVAVYVEFCYKGNLSNFRKARDASPCAEKSPQGRGF 960
 QY 961 RAMVELARLDRRRPGSSDRVLFAFRTKTEGGARRASPDQAEADLWLSPLTMEDLVCSYFQ 1020
 DB 961 RAMVELARLDRRRPGSSDRVLFAFRTKTEGGARRASPDQAEADLWLSPLTMEDLVCSYFQ 1020
 QY 1021 VARGMEFLASRCKIHRDLAARNILLSSEDDVYKICDFGLARDIYKDPVYRKGSRARLPKW 1080
 DB 1021 VARGMEFLASRCKIHRDLAARNILLSSEDDVYKICDFGLARDIYKDPVYRKGSRARLPKW 1080
 QY 1081 MAPEISFDKVTYTSQDVNSFVLLWEIFSLGASPPGVQINEEFQRLDGTMRAPELA 1140
 DB 1081 MAPEISFDKVTYTSQDVNSFVLLWEIFSLGASPPGVQINEEFQRLDGTMRAPELA 1140
 QY 1141 TPAIRIMNCWSDGPKAPAFSELVEILGDLQGRGLQEEVCMAPRSOSSSEGSFS 1200
 DB 1141 TPAIRIMNCWSDGPKAPAFSELVEILGDLQGRGLQEEVCMAPRSOSSSEGSFS 1200

Tue Jan 23 09:28:37 2001

821 EWPLEECEYLSDASQWERPRELHLGVLGYGAFGKVVVEASAFGIHKGSSCDTVAVKM 880
 DB 801 EYVIDEGERLPYDASQWEPFDRDLKGLKPLGAGFGQVIEADAFGIDTACTRTVAVKM 860
 QY 881 LKEGATASEHRALMSELKILIHIGHNLNVNLLGACTKPGQPLMWIVEFCKYGNLSNFLR 940
 DB 861 LKEGATHSEHRALMSELKILIHIGHNLNVNLLGACTKPGQPLMWIVEFCKYGNLSAYLR 920
 QY 941 AKRDAFSCAKSPEQGRFRFAMVELARLD-----RRPGSSDRVLFAFESKTEGGARRA 995
 DB 921 SKRSEFIPYRKVS-----ARFQGRKENTGDLSTDLKOR---LDSITSSQSSTSSGGVEER 973
 QY 996 S-----PDQAEADLWLSPLTMEDLVCSYFQVARGMEFLASRKCIHRDLAARNILLSSED 1049
 DB 974 SILSDVEEDAGSLEDCKNPLTMEDLVCSYFQVARGMEFLASRKCIHRDLAARNILLSDDN 1033
 QY 1050 VVKICDFGLARDIYKDPYVYKGRARPLKWMAPESIFDKVYITQSDVWSFGVLLWEIFS 1109
 DB 1034 VVKICDFGLARDIYKDPYVYKGRARPLKWMAPETIFDRVYITQSDVWSFGVLLWEIFS 1093
 QY 1110 LGASPYPGVQINEEFQCRILRDTGRAPELATPAIRIMLNCWSGDPKARPAFSELVEIL 1169
 DB 1094 LGASPYPGVKIDEEFCRLKEGTRMRAPDYTTPEMYOTMLDCHWDGPKQRPTESELVEHL 1153
 QY 1170 GLLQGRGLQEEVECMAPRS-SQSSEGSFSQVSTMALHIAQADAEDSPSLQRHSLAA 1228
 DB 1154 GNLLQANVRQDGKDYVVLPSVSLNNEEDSGLSUPTSPASKEEEFVCDPKFYHDNTAGI 1213
 QY 1229 RYNNVWSFPFGCLARGAETRGSSR-----MKTFFEEPMTPTYTKGSDVNDTDSGMWLASE 1282
 DB 1214 SQYR-----QGSKRSRPSVSKTFEDIPLVTVTVVQVEENOTDSDGMWLASE 1259
 QY 1283 EFQIESRHRQESGSCKPGQGNQAVTAHDPDSQGRRRRPERGARGGVFNSYEGELSE 1342
 DB 1260 ELKTLQEDQK-----VKIPFTSLAPSKSNESVMEASNOTSGYQSG---YHSDMDMNAV 1311
 QY 1343 PSEEDH---CSPSARVT 1356
 DB 1312 CSSEDTYELCAQEAAPT 1328
 RESULT 5
 S51656
 C:Species: vascular endothelial factor receptor - quail
 C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999
 C:Accession: S51656
 R:Eichmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.
 A:Description: submitted to the EMBL data Library, December 1994
 A:Reference number: S51656
 A:Accession: S51656
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1348 <EIC>
 A:Cross-references: EMBL:X83288; NID:g603523; PIDN:CAA58268.1; PID:g603524
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
 C:Keywords: ATP
 F:823-1160/Domain: protein kinase homology <KIN>
 F:831-839/Region: protein kinase ATP-binding motif

Query Match	38.9%;	Score 2804;	DB 2;	Length 1348;
Best Local Similarity	43.9%;	Pred. No. 2.8e-121;		
Matches 613:	Conservative 218;	Mismatches 456;	Indels 110;	Gaps 31;

1 MORGAAALCLRLWLCGL-LLDGLVSGYSMTPTPLNTTESHVYIDTGDSLSLSCRGQHPLEW 59
:
:
:
1 MELGPRVLTVLLCLAPVFAGLF--ISMDPPLSLQKSVLTITTTNDTLNITSRGRAYVV 58
:
:
:
60 AWPGEAQAPATGDKDSEDGTGVWRDC-EGTDARPYCKVLLLHEVIANDTGSVCYYKYIKA 118
:
:
:

59	SWPNQSS-----VEKRLAVTGCSEG-----PFCKTLLRIVGNDTGDRCILYG----	101
119	RIEGTTAASSVYVRFQEPFINKPOTLL-----VNRKDAHWVPCULVSIGLNVLRSQ-----	172
104	--DSOARTTIVYVQDYSRPFVTSVGDLGVIYITKNKTIVVVPVLCGTIVNLNVSLRAKYP	161
173	SSVLWPGQGVWDDRGMLVSTPLLDALYLQCEITWGDQDFLSNPFLVHITGHELYDI	232
162	EKFVDPGKISWDNKGGFTIPSHLYNAGHWFECEAKIDNESYSQYIIVAVVGRIYDL	221
233	QLLPKRSLELLVGEKVLNCTVWAEFNSGVTFDWDYPGKOABRGKWVPERRS-----OOT	287
222	TMNPHYQVELAVEKGLVNLCTVTELVNGIDFRWDYPS-----IKERRATIRDLKTT	273
288	HTELS-----SILTIHNVSQHDLSGVYCKCANNGIORFRETEVIVHENPFFISVEMLGP	344
274	AGEIKTFVSTLTIESVNLSDKGRVTCASSGRNMKNSSYFIHESPFILHEKME--VYE	332
345	ATAGDELKVLKLAAYPPPEFOFYKDGAALSGRHS---PHALVKEVTEASTGYTIAL	401
333	MKLGD-TVSIIPVKFKGTPPPKAKYKNGKVINANHTVTKLYALVITEATEKDAGNTVVL	391
402	WNSAAGLRNRISLELVNVPVPIHEKPAESP-SIYSRHSQALCTAYGVPLPLSTQWHE	460
392	TNPTNMQMKRUTFTLLVNVVPOIGENALMAPVDSYKYSTQALCTCTIYAVPPAAVLMTN	451
461	RPTWCKMFAQBSLRROQOQDLMPQCDWRNAVITQDAVNPESLDTWTEFEVEKKNVTSK	520
452	QLEEEC-TFSPOKVRLGANP---YACRKWKVISEKGGNQVE-IKQRVYTIAGKTIVST	506
521	LVIQNANTSAMTKCVSVKNVQDERLIYFYVTTIPDGTFTIESKPSBELLEGQPVLLSCOA	580
507	LVIOAANYALRYCATNTRAGSSSERVISFHVTR---GLEINLQPSQLTEKMDTSLQCTA	563
581	DSYKTEHLRWLRNLSTLHDAGHPLLLDCKNVHLFATPLAASIEEYAPGARATLSLSI	640
564	DKTFEKLWSYKLSHVTSQTFPGPLPMPVCKNDAL-OKLNATVSNV--NGENVTLIELIL	620
641	PRVAPHEGHYVCEQDRRSHDKCHKYLSVQALEAPRLTNLDLLVNVSDLEMKQCL	700
621	RNISLDQGDYVCIQAQDKAKTQCHLVKHLTVQEPHLRVLNGLENQTTNIGETIEVLCT	680
701	VAGAHAPSTVWTKDERLLEEKSGVDLADSNOKLSIQRVREDAGRYLCVNCNAGCVNSS	760
681	VNGVPPNITWFKNSETLFDSSGIVLKDGKNTLIRRVKREDGGLYTCLACNLTIGCKAAE	740
761	ASVAVGSGEDKSMETILVGTGVTAPEFWLLLLIFCNMRPRADIKTYLSIIMDPG	820
741	AFSVQAEKTNLEILVLGTVATVMEFWLLVILLETVRANGDMKTYLSIIMDPD	800
821	EYPLEQCEYLSYDASOWEFPFRERUHLGRVLGYGAFGKVEASAFGHKSGSCDTVAVKM	880
801	EVPIDECHERLPYDASKWEEFPDRLLKGLPGRGAFGQVIEADAFGDKATCTRTVAVKM	860
881	LKEGATASEHRALMELKTLIHIGNHANVNLGACTPOGPLAVIVFECKYGNLSNFLR	940
861	LKEGATSHSRALMSLKTLIHIGHNLNVNLLGACTAPGGPLAVIVFEYKFGNLSAYLR	920
941	AKRDAFSPCAEKSPEQGRFRANVELARLD-----RRRPGSSDRVLPARFSTKEGGARRA	990
921	SKSEIPIYKMKSS---AREFGOKENYTGDISDCLKQR---LDSITSSQSSTSGFVEER	970
996	S-----POEAEIDLWSPLTWEDIVCVYSFQVARGHEFLASRKCIHRDLAARNILLS	1040
974	SLSDVFEEDAGSDILCKNPLTMDILICYSFQVARGHEFLASRKCIHRDLAARNILLS	1030
1050	VVKICDFGLARDIYKDPDYVRKGSARLPLKMWAPESIEDKYVTTQSDVMSFGVLLWEIFS	1110
1034	VVKICDFGLARDIYKDPDYVRKGDARLPLKMAPETIFDRVYTIQSDVMSFGVLLWEIFS	1070
1110	LGASPYPGVOINEEFCORLUDGTFMRAPELATPATRIMLCNCSODKPARPAFSELVEIL	1170
1094	LGASPYPGVIDEFCRRLKEGTMRAPDYTPPYMTQMLDCHWDGDKORFTSELVEHL	1130

QY 1170 GDLQGRGLQEEVEECMAPRS-SQSEEGSFQVSTMLHIAQADAEDSPPSLQRHSLAA 1228
DB 1154 GNLQANVRQDGDYVVLPLSVLNMEEDSGLSLTPASCSAEVEECVDPKPHYDNTAGI 1213
QY 1229 RYNNWVSFPCLARGAETRGSSR-----MKTFFEPNPTTYIKGSDVNDQDSGHVLAASE 1282
DB 1214 SQYR-----QGSRRKSRPVSVKTFEDIPLVTVVQVEENQDQDSGHVLAASE 1259
QY 1283 EFOQIESHRQESGFSCKPGQVAVTRAHPDSQGRRRRPERGARGGQVYFNSEYGEASE 1342
DB 1260 ELKLEBQDQK-----VKIPFSTLAFSKSNESVMSEASNOTSGYQSG---YHSDMDNMV 1311
QY 1343 PSEEDH---CSPARVY 1356
DB 1312 CSSEDTLLCAQEAASPT 1328

RESULT 6
JC1402
Protein-tyrosine kinase (EC 2.7.1.112) KDR - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 28-Aug-1998 #text_change 04-Feb-2000
C:Accession: JC1402; I58357
R:Termin, B.I.: Dougher-Vermazen, M.; Carrión, M.E.; Dimitrov, D.; Armellino, D.C.; Gosh
Biochem. Biophys. Res. Commun. 187, 1579-1586, 1992
A:Title: Identification of the KDR tyrosine kinase as a receptor for vascular endothelial
A:Reference number: JC1402; MUID:93038639
A:Accession: JC1402
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-48, 'I', 50-616, 'E', 618-700, 'E', 702-748, 'F', 750-1263, 'I', 1265-1290, 'P', 1292-
A:Cross-references: EMBL:X61656; NID:g31717
R:Termin, B.I.: Carrión, M.E.; Kovacs, E.; Rasmussen, B.A.; Eddy, R.L.; Shows, T.B.
Oncogene 6, 1677-1683, 1991
A:Title: Identification of a new endothelial cell growth factor receptor tyrosine kinase
A:Reference number: 158357; MUID:92019839
A:Accession: I58357
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 3-1356 <TER>
A:Cross-references: GB:I04947; NID:g186674; PID:AAA59459.1; PID:g186675
C:Comment: This enzyme acts as a receptor for vascular endothelial cell growth factor.
C:Genetics:
A:Gene: GDB:KDR
A:Cross-references: GDB:127921; OMIM:191306
A:Map position: 4q12-4q12
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein ki
F:766-790/Domain: transmembrane #status predicted <TM>
F:832-1167/Domain: protein kinase homology <KIN>
F:840-848/Region: protein kinase ATP-binding motif

Query Match 38.2%; Score 2753; DB 2; Length 1356;
Best Local Similarity 43.9%; Pred. No. 6.2e-119;
Matches 606; Conservative 205; Mismatches 480; Indels 88; Gaps 26;

QY 1 MORGALRLWLGLGLDGLVS--GYSWPTPLNITEESHVIDTDSLSISCRQHPLE 58
DB 1 MESKVLVALVLCVETRAASVGLPSVLDLFRUSIQKDLITIKANTTLQITCRGQNDL 60
QY 59 WAPGAQAPATGDKDSEDTGVVRC-EGTDARYPKYKLLHVEHNDTGSVYCYRYIK 117
DB 61 WLWPNQSG-----SEORVEVTECSDEL----FCKTLTIPKIGNDTGAYKCFYR--- 106
QY 118 ARIEGTATASSVYVRDEQPIPKPDT-----LLVNRKAMWVCLVSLPGLNVIL-- 169
DB 107 ---ETDLASVYIVQDVRSPFIASVSDQGVVYITENKNTVWIPCLGSLNLSVLSA 163
QY 170 RSQSSVLPDQGVVMDRRRGLMVTPLHDAALYLCQETTWGDDFLSNPFLVHTIGNEL 229
DB 164 RYPEKRFVDPGNRISWDSKGTIPSYMSYAGVFECAKINDESYQSIMYIVVGVYRI 223

QY 230 YDIOLLPRKSLLELVGKLVNCTVWAEFNSGVTFDWDYPGKOAERGKWYPERRSQOTHT 289
DB 224 YDVLSPSHGIELSVGKLVNCTARTELNVDINFEYFSSKHQKLVNRLDKTQSGS 283
QY 290 ELS---SILTIHNVSDHLAGSYVCANNQIORFRESTEIVHNPFFISVEMWLGPILEAT 346
DB 284 EMKFLSLTIIDGVTSDQGLYTCAASSGLMTKKNSTFVRVHEKPPFAFGSGMESLYEAT 343
QY 347 AGDELVLKLVKLAAPPPEQWYKDGKALSQRH---SPHALYKLVKEVTEASTGYTLALWN 403
DB 344 VG-ERVRIAPKLYGYPPEIKWYKNGIPLESNHTYKAGHVLTIMEVSESDTGNVTILTN 402
QY 404 SAAGLRNLSLELVVNPVPPQIHEKEASSP-SIYSRHSRQALTCCTAYGVPLPLSQHWHP 462
DB 403 PISKEQSHVSVLVVYVPPQIGEKSLSPVDSYQYGTITLTCTVYAIPPPHEHMTWOL 462
QY 463 WTPCKMFAQSLRRROODLMPQCDRHRVTTQDANPIESLDTWTFEVEGKNKTVSKLV 522
DB 463 EEEC---ANEPQAVSVTNYP-CEENRSVEDFOGKNKIEYKNKNOFALIEGKNKTVSLV 518
QY 523 IQANVNSAMYKCVNSKYGQDERLIYFYVTTIPDQFTIESKPSBELLEGQPVLLSCQADS 582
DB 519 IQANVNSALYKEAVNKGVRGERSVIFVTRGPE---ITLQPDMDQPTQESVSLWCTADR 575
QY 583 YKYEHLRWYRLNLSLTHDAHGNPLLLDCKNV----HLFATPLAASLEAVAPGARHATLSL 638
DB 576 STFENLTWYKLGPOPLPIHVGELEPTPVCKNLTDLWKNLMTFMSNSTNDI-----LIM 627
QY 639 SIPRVAPEHEGHYVCEVDRRSHDKCHKVLSVQALCAPRLTONLTLLVNVSDSLMQ 698
DB 628 ELKNASLODQDGYVCLAQDKTKRHCVVRLTVLERYAPTITNLENOTTISIGESIEVS 687
QY 699 CLVAGAHAPSIVWYKDERLEEKESQVLDASQKLSIORVREEDAGRYLCSVCNAGCVN 758
DB 688 CTASGNPPQIMWFEDNETLVEDSGIVLKDGNRLTIRVRKDEGLTYCQACSVLGCAK 747
QY 759 SSASVAVESGDKGSMETVILVGTGVIAVFFWVLLLLIFCNRRRAHAADIKTYLSIMD 818
DB 748 VZAFIIEGAQEKTNLEIILVGTIVIAEFWVLLVILGTVKANGGELATGYLSIMD 807
QY 819 PGVEPLEEOCYLSYDASOMEPPRELHLGRVLYGAFGVKVEASAFGIHKGSCDVTAV 878
DB 808 PDELPLDEHCERLPYDASKWEPPRDLNLGPKLGRGAFQGEIADFADGIDTATCTAV 867
QY 879 KMLKCATASEHRALMSCLKILIHGNHNVNVLGACTKPGPLMVIVFCYKGNLSNF 938
DB 868 KMLKCATASEHRALMSCLKILIHGHVNLVNLGACTKPGPLMVIVFCYKGNLSNF 927
QY 939 LRAKRDASFCAEKSPEQGRFRA---MVYELARLDRRRRFGSSDRVLFARFSKTEGGARRA 995
DB 928 LRSKRNEFVYPTKG---ARFRQGDYVGAIPVDLKR--RLDSITSSQSSASSGFVEEK 981
QY 996 S-----PDQEA-EDLWLSPLTMEVLVCYSQVARGHEFFLASRCKIHRDLAARNILLSSEV 1050
DB 982 SLSDVEEEEPEDLYKDFLTLEHLCYSFQVAKGMEFFLASRCKIHRDLAARNILLSSEV 1041
QY 1051 VKICDFGLARDYKDPDYVRKGSARLPKWMAPESIFDKVYTTQSDVWSFGVLLMEIFSL 1110
DB 1042 VKICDFGLARDYKDPDYVRKGDARLPKWMAPETIFDRVYTTQSDVWSFGVLLMEIFSL 1101
QY 1111 GASPYGVQINEEFCORLRDGTMRAPELATPAIRIMHNLCSGDKPAPAPSELVEILG 1170
DB 1102 GASPYGVQIDEFFRLKEGTRMRAPDPTTPMYQTHLDCWHGEPQSPTSELVEHILG 1161
QY 1171 DLQGRGLGEEEEVCNAPRS-SQSEEGSFQVSTMALHIAQADAEDSPPSLQRHSLAA 1229
DB 1162 NLQANQDQKDYIVLPISETLSMEEDSGLSLTPSPVSCMEEEVCDP-----KPH 1213
QY 1230 YNNVSPFCCLARGAETRGSSRMKTFEPPM-TPTTYKGSVDNQTDSGHVLAASEFEQIE 1288
DB 1214 YDNTAGISQVYLRNKRKSRPVSVKTFEDIPLEEPKVIIPDDNQTDSGHVLAASEELKTL 1273

QY 1289 SRHROESGCKPGQNVAVTRAHPDSQGRRRRPERGARGQGVFNSEYGESEED 1347
Db 1274 DTKLSPSGCMVPSK-----SRESVASEGNSQ-----TSYGQSGYSDTDITVYSSEE 1323

RESULT 7
A11228
protein-tyrosine kinase (EC 2.7.1.112) Flk-1 precursor, endothelial cell-specific receptor
C:Species: Mus musculus (house mouse)
C:Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 04-Feb-2000
C:Accession: A41228; A46065; I58365; S18832; S29991
R:Matthews, W.; Jordan, C.T.; Gavin, M.; Jenkins, N.A.; Copeland, N.G.; Lemischka, I.R.
Proc. Natl. Acad. Sci. U.S.A. 88, 9026-9030, 1991
A:Title: A receptor tyrosine kinase cDNA isolated from a population of enriched primitive
A:Reference number: A41228; MUID:92020984
A:Accession: A41228
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1367 <MAY>
R:Cross-references: GB:X59397; NID:g50976; PIDN:CAA42040.1; PID:g50977
Cell 72, 835-846, 1993
R:Willauer, B.; Witzmann-Voos, S.; Schnurch, H.; Martinez, R.; Moller, N.P.; Risau, W.
Cell 72, 835-846, 1993
A:Title: High affinity VEGF binding and developmental expression suggest Flk-1 as a major
A:Reference number: A46065; MUID:93208880
A:Accession: A46065
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-24, 'T', 26-782, 'VL', 785-916, 'C', 918-1367 <MTL>
A:Cross-references: GB:T08842; NID:g57923; PIDN:CAA50192.1; PID:g57924
A:Note: submitted to the EMBL Data Library, January 1993
R:Oelrichs, R.B.; Reid, H.H.; Bernard, O.; Ziemielski, A.; Wilks, A.F.
Oncogene 8, 11-18, 1993
A:Title: NTK/FLK-1: a putative receptor protein tyrosine kinase isolated from E10 embryonic
A:Reference number: I58365; MUID:93141255
A:Accession: I58365
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-678, 'D', 680-1340, 'RSPV' <OEL>
A:Cross-references: GB:S53103; NID:g264004; PIDN:AAB25043.1; PID:g264005
C:Genetics:
A:Gene: FLK-1; NTK
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; transmembrane
F:830-1165/Domain: protein kinase homology <KIN>
F:838-846/Region: protein kinase ATP-binding motif

Query Match 37.08; Score 2669; DB 2; Length 1367;
Best Local Similarity 43.08; Pred. No. 4.4e-115;
Matches 588; Conservative 205; Mismatches 485; Indels 90; Gaps 24;

QY 1 MORGAALRLWLCGLLDGLVS--GYSMPTPTLNITEESHVIDTGDLSISCRGHPLE 58
Db 1 MESALLAVALWFCVETRAASVGLPGLPPLKSTQKDLTILANTLQITCRGQRLD 60

QY 59 WAWPAQAPATGDKDSEDGTGVRDCEGTDAEPYCKVLLHREHANDTGYCYRYIKA 118
Db 61 WLWPAQ-----RDSEERVLVTECGGDS-IFCKLTIPRVVNDTGAYKCYRDV-- 110

QY 119 RIEGTTAASSVVFVRDEQPIKPDIT-----LLVNRKAMWVCLVSLPGLNVL--R 170
Db 111 ----DIATVVTVYRDRSPFIASVSOHGIYVITENKNTVPIPCRGSIINLVSLCAR 166

QY 171 SQSSVLWPDGVVWDDRRGLVSTPLLDHALYLQCETTWGDQDFLSPFLVHTGELY 230
Db 167 YPERKFPDGNRISWSEIGFTLPYSYMSVAGWVCEAKINDEYQSIWIVVVGWY 226

QY 231 DIQLPKRSELLGKVLNCTVWAEFNSGVTFDWDYPCQKRGKWKPERSSQOHTHE 290
Db 227 DVLSPPEHELSAGEKLVNLCTARTNELNVLGDTWHSPPSKSHHKIVNRDVKPFGTV 286

QY 291 LS----SILTIHNVSDHDLGVSVCANNGCIQRFRESTEIVHENPFIISVWLKGPILATA 347

Db 287 AKMELSTLTIESVTKSDQGEYTCVASSGRMIKRNRTFVRVHTKPFIAFGSGMKSLVEATV 346
QY 348 GDELVKLPVKLAAYPPPEFQWTKDKALSGRHS--PHALVKEVTEASTGYITLALWNS 404
Db 347 GSO-VRIPIVKYLSYAPADIKYRNGRPIESNTYMTVGDELTIWETDAGNYTILNTP 405
QY 405 AAGLRNLSLELVNVPPOIHEKEASSP-SIYSRHSROALCTAYGVPLSLQIOWHWP 463
Db 406 ISMEKQSHVSLVNVNPPPOIGERKALISPMDSYQYTMQLTCTVYANPPLHHIOWTQLE 465
QY 464 TPKFAQRSLRRQQDLMPQCRDWRVAVTTQDAVNPIESLDTWTFEVEGRKNTYSLVY 523
Db 466 EACSY-----RPGQTSYACKEMWRHVEDFOGKNKIEVKNOYALIEGKNKTSTLVI 517
QY 524 QNANYSAMYKCVSNKVGQDERLIIFYVTTPDGFTIESKPEEELLEGOPVLLSQADSY 583
Db 518 QAAVNSALYKCEAIKAGRGERSVIFHVIRGPE---ITVQPAAQTEQESVSLCTADR 574
QY 584 KYEHLRWYRLNLSLTHDAHGNPLLLDCKNV----HLFATPLAASLEEVAPGARHATLSL 639
Db 575 TEENLTWYKLGSOATSVHGESLTPVCKNLDALWKLNGTMSNSTNDI-----LIYA 626
QY 640 IPRVAPHEGHYVCEVQDRRSHDKKYLVSQALAPRLTONLTDLLNVNVSLSLEMC 699
Db 627 FONASLDQGDYVCSAODKKTKRKHCLVQLIILERMAMPITGNLENQTTIGETIEVTC 686
QY 700 LVGAHAPSIVMYKBERLEESGVDLADSNOKLSIQRVREEDAGRYLSCVNAKCVNS 759
Db 687 PASGNPTPHITWFKDNETLVEDSGVLRDGNRNLTIRRVKEDGGGLTYCOACNVLGACA 746
QY 760 SASVAVEGSEDGSMIEVILGTGVIAVFFWVLLLIIFCMRRPAPAHADIKTGYLSINDP 819
Db 747 ETLFIEGAQKTNLEVIILGTAVIANEFLLVLLVILVTRKANEGLKTYGLSIVMDP 806
QY 820 GEVPLEEQEYLSYDASQWEPFRERLHGRVLYGAGFVWZASAFHKGSCDTPVAK 879
Db 807 DELPIDERCERLPYDASKWEFPRDLKLGKPLGRGAFQVIEADAFDKTATCTKTVAK 866
QY 880 MLKEGATASEHRLMSALKILTHIGNHNVNLLGACTKPGGLVIVIFCYGNLSNL 939
Db 867 MLKEGATASEHRLMSALKILTHIGHHNVNLLGACTKPGGLVIVIFCYGNLSNL 926
QY 940 RAKRDAFSAEKSPQGRFRA----MVELARLDRRRFGSSDRVLFARFKTEGGARRA 995
Db 927 RGRNREFVYKSGK-----ARFROGKDYVGLS-VDLKR--RLDSITSQSSASSGPFVEK 979
QY 996 S-----PDQAEEDLWLSPLTMDLVCSYQVARGMEFLASRKCIHRDLAARNILLSSE 1050
Db 980 SLSDVEEESEELYKDFLTLEHLICYSFOVAKGMEFLASRKCIHRDLAARNILLSSE 1039
QY 1051 VKICDFGLARDYKDPDYVRKGSARLPKWPAPESIFDKVYTTQSDVWSFGVLLWEIFSL 1110
Db 1040 VKICDFGLARDYKDPDYVRKGDARLPKWPAPETIFDRVYTIQSDVWSFGVLLWEIFSL 1099
QY 1111 GASPYGVOINEEFQORLDRGTNRAPELATPAIRIMLNCWSDGPKAPAFSELVEILG 1170
Db 1100 GASPYGVIDEFCRLEKGTNRAPADYTPMYQTMDCWHDHPNQRPSFSELVEHLG 1159
QY 1171 DLLOGRGLQEEVEVCNAPS-SOSSSEGSQSVSTWALHIAQADAEDSPSLORHSLAR 1229
Db 1160 NLLQANQDQGDYVILPMSETLSMEEDSGLSLTPSVSCHEEEVCDP-----KPH 1211
QY 1230 YNWNVFPQCLARGAETRGSRMKTTEEPFM-TPTTYKGSVDNQTDSGMVLAEEFEQIE 1288
Db 1212 YDNTAGISHVLQNSKRKSPVSKTFEDIPLEPEVKVIPDDQSDTDSGMVLAASELKTLE 1271
QY 1289 SRHQESGSCPKPGQNVAVTRAHPDSQGRRRRPERGARGQGVFNSE 1336
Db 1272 DRNKLSPSFG----GMHPSKSRSVASEGNSQ-----TSYGQSGYHSD 1310

RESULT 8

I78875
receptor tyrosine kinase - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 24-Sep-1999
C:Accession: I78875
R:Finerty, H.; Kelleher, K.; Morris, G.E.; Bean, K.; Merberg, D.M.; Kriz, R.; Morris, J.
Oncogene 8, 2293-2298, 1993
A:Title: Molecular cloning of murine FLT and FLT4.
A:Reference number: I58375; MUID:93330572
A:Accession: I78875
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1333 <RES>
A:Cross-references: GB:L07297; NID:G293782; PID:AAA40078.1; PID:G293783
C:Genetics:
A:Gene: FLT
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C:Keywords: ATP
F:826-1161/Domain: protein kinase homology <KIN>
F:834-842/Region: protein kinase ATP-binding motif

Query Match 34.2%; Score 2469; DB 2; Length 1333;
Best Local Similarity 42.4%; Pred. No. 6.3e-106;
Matches 555; Conservative 186; Mismatches 472; Indels 96; Gaps 27;

QY 5 AALCLRLMCLGLDGLVSGYSMTPTNTITEESHVIDTGDLSLSCRGHPLKAWPGA 64
DB 8 AVLPLVALLGCL-LRTGYSGSLKLPVLSLKGTHQVMAGQTLFLKCGEAHNSLP-- 64
QY 65 QEAPATGDKSDTGVWRDCEGTARPCKVLLLEHVEHNDTGSYVCYKIKARIEGTT 124
DB 65 ---TTVQEDRKLSTPSSACGRDNQFCSTLTDTAQAHTGLVTCRYLPTSTSKRKA 121
QY 125 AASVYVVRDFQPIPK-----PDLLVNRKDMVPLVSLPGLNLTLS-QSSVLWD 179
DB 122 ESSIIYFVSDAGSPFIEMHTDIPKLVHTEGROLIIPCVSPNTVTLKFPEDTLTPD 181
QY 180 GOEYVVDORGLMVLPTLLHDALYLQCETTMGDQDFLNPFLVHTGELNYDIQLLPKS 239
DB 182 GORIWDSSRGFIINAYNYKEIGLNCENTVNGHLYQIN-YLTHRQNTILVDQIRPSP 240
QY 240 LELLVGEKLVLNCTVNAEFGNSVTDWDPYQKQKRWPER--RSQOHTFELSLTI 297
DB 241 VRLHGGQTLVNLCTATTELNTRVQSNYPGKATKAS-IRQIDRSHNNVFSVLKI 299
QY 298 HNVSOHDLGSYCKKANGIORFRESPEVIVHENPFTSVMLKGPILFATAGDELVKLPVK 357
DB 300 NVNESDKGLITCRVKSQSFQSFNTSVHYEKGFIKVPQVQETTAGRSYRLSK 359
QY 358 LAAYPPPEFQWKDGLKAL---SGRHSF--ALVKEVTEASTGTYTLALNWSAAGLRNI 412
DB 360 VKAFPSPEIWLKDGSPATLKSARYLVHGSYLIKDVTEDAGDYTLILGIRKSLRNL 419
QY 413 SLELVNVPQIIEKEAS---SPSIYSRHSQALCTAYGVLPLSIOHWHRPWTCKMF 469
DB 420 TATLVNVPQIIEKESVSSLPPLPLGSRQVLTCTVIGIPRP-TITLWH---PC--- 472
QY 470 AQRSLRRKQDLMPCQRD-WRAVITQDAVNPVIESLDTWTEFVEGKNKTVSLQIONAVY 528
DB 473 --HNHSKERYDCTENESEFILDPSNGLNRESISQRMVTEGNTKVTSLVAVDSQT 530
QY 529 SAMYKCVSNKVGQDERLIYFTYTPDGF--TIESKPEELLEGOPVLLSCQADSYKE 586
DB 531 PGISYCRAFNKIGTVERNIKFYVTDVNGFVHVSLEKMPA---EGEDLKSCVNVNFKLR 586
QY 587 HLRWYRL---NLSTLHDAGHNPILLCKNVHLFATPLAASLEEVAPGARHATLSIPRV 643
DB 587 DITWILLRTVNNRTMH-----HSISKQKMATQDYS-----ITLNLVKNV 627
QY 644 APEEGHYVEVQDRSHDKCHKYLSVQALEAPRLTQNTDLLNVNDSLEMOCLVAG 703
DB 628 SLEDSGTACRARNIYTGEDILTRKTEVLVRDSEAPHLLQNLSDYEVSGSTTLDQARG 687

QY 704 AHAPSIVWTKDERLLEEKSGVDLADSNOKLSIORVREEDAGRYLCSYNAKGVNSSASY 763
DB 688 VPAQITWFKNNHKOQEPGIIILPGCNSTLFIERTVEDEGVYRCRATNOKGAVESAAYL 747
QY 764 AVEGSEDKSGMEIVILGVGVIAVFWVLLLLIFCNMRPAPAHADIKTYLSIIMDPGEVP 823
DB 748 TVQGISDKSNLELITLTCTCAATLFWLLTLFIRKLKR-SSSEVKTYLSIIMDPGEVP 806
QY 824 LEEQCEYLSDASQWEPFRERHLGRVGLGAGFKVVEASAGSIHKGSCDDVAVAKLKE 883
DB 807 LDEQCERLPTDASKWFEARELRKLKSLGAGFKVQVQASAFGIKKSPICRTVAVAKLKE 866
QY 884 GATASCHRALMSELKILIHGNHNVNLLGACTKPGQLMIVVEFCYKGNLSNLRKR 943
DB 867 GATASEYKALMTELKILTHIGHNLVNVLLGACTKGGPLMIVVEYCKYKGNLSNLSKR 926
QY 944 DAFSPCAEKSP-----QRGRFAMVELARLDRRRGSSDRVLFARF-----SKTEGGA 992
DB 927 DLF--CLNKDAALHMLKESLEPGLEQOKPRLDSVSSSVTSSFPEDRSVSDVEG-- 982
QY 993 RRASPOEAEEDLWLSPLTMEDLVYCSFOVARGMEFLASRKCIRDLAARNILLSESDEVK 1052
DB 983 -----DEDYSEISKQPLTMEDLISYFQVARGMEFLSSRKCIRDLAARNILLSENNVK 1037
QY 1053 ICDGLARDIYKDPDYVRKGSARLPLKMAPESIFDKVYTTQSDVWSFGVLLWEIFSIGA 1112
DB 1038 ICDGLARDIYKPNPDYVRGDTRLPLKMAPESIFDKVYTKSDVWSYGVLLWEIFSIGG 1097
QY 1113 SPYGVQINEEFQORLDGTRMAPELATPAIRMLNCWSDGPKARPAPFSELVEILGDL 1172
DB 1098 SPYGVQMDQDFCSRLKEGMRMTPEYATPEIYQIMLDCWHKDKPRPAELVEKLGDL 1157
QY 1173 LQGRGLQEEVEVCMAPRSSQSSSEGSFVSQVSTMALHTAQADAEDSPSLQHSLLAAYTN 1232
DB 1158 LQANVQDGRD--YIPLNALITRNSSTYSTPTFSEDLFKDGFADPHFSGSSDDVRYN 1215
QY 1233 WVSFPGCLARGAETGRSSRMKTFFEPFMTTYKSGVDNQTDSGMVLAS 1281
DB 1216 AKRF-----MSLERIKTFEE--LSPNSTSMFEDYQIDYSTLLGS 1252

RESULT 9
S09982
protein-tyrosine kinase (EC 2.7.1.112) f1t1 precursor - human
N:Alternate names: receptor-type tyrosine kinase fit
C:Species: Homo sapiens (man)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 20-Jun-2000
C:Accession: S09982
R:Shibuya, M.; Yamaguchi, S.; Yamane, A.; Ikeda, T.; Tojo, A.; Matsushima, H.; Sato, O.
Oncogene 5, 519-524, 1990
A:Title: Nucleotide sequence and expression of a novel human receptor-type tyrosine k
A:Reference number: S09982; MUID:90221591
A:Accession: S09982
A:Molecule type: mRNA
A:Residues: 1-1338 <SHI>
A:Cross-references: EMBL:X51602; NID:G31431; PIDN:CAA35946.1; PID:G31432
C:Genetics:
A:Gene: GDB:FL11
A:Cross-references: GDB:120616; OMTM:165070
A:Map position: 13q12-13q12
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP; autophosphorylation; glycoprotein; membrane protein; phosphoprotein;
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-1338/Product: protein-tyrosine kinase fit #status predicted <MAT>
F:825-1161/Domain: protein kinase homology <KIN>
F:833-841/Region: protein kinase ATP-binding motif

Query Match 34.2%; Score 2463; DB 2; Length 1338;
Best Local Similarity 39.9%; Pred. No. 1.2e-105;
Matches 547; Conservative 228; Mismatches 469; Indels 128; Gaps 28;

QY 4 GAALCLRLWLCIGLDGLVSGYSMTPTLNTITEESHVIDTGDLSISCRGQHPLEAWPG 63
Db 8 GVLLC-ALLSCL-LLTGSSGSKLDPELSLKGTHQINQAGOTLHOCRGAAHKS---62
QY 64 AQEAPATGDKDSEDGVVYDCGTDPKCKVLLHVEHANDTGSVYCYIKARIEGT 123
Db 63 ---LPEWSEKESRUSITKACGRNGKQFCSTLTNTAQANHTGYSKYLAVPSTKKKE 119
QY 124 TAASSVVFYRDFEQPFINK---PDTLLVNRKDMVWVCLVSLPGLNVTLS-OSSVLPW 178
Db 120 TESAIYIFISGRPEVEMYSIPEIHTMEGRELVIPCVTSPTNITVLAKFPLDTLP 179
QY 179 DGEVVDORRGMVSTPLLDALYLOCTETWGDQDFLSNPLVHTGELNLYDIOLLPK 238
Db 180 DGRRIIWDGRKFIISNATYKEIGLTCATVNGHLYKTN-YLTHROTNTIIDVOISPTR 238
QY 239 SLELLVGEKLVNCTVMAEFNSGVTPDWYPGKAERGKWPERRSQOHTELS---SIL 295
Db 239 PVKLLRGHTLVNCTATTPLNTRVQMTWSPDEKNRAS--VRRIDQNSHANIFYSVL 296
QY 296 TIHNSHDLGVSVCNANGIQRFESTEVIVHENFISVEMLKGPILFATAGDELKLP 355
Db 297 TIDKMQKRGYTCVRSGPSEKSVNTSVHIYDKAFITVKHKKQVLETVAGKRSYRLS 356
QY 356 VKLAAYPPPEFQYKDGKALSGRHS-----PHALVKEVTEASTGYTTLALMNSAAGLR 410
Db 357 MKVAFSPPEVWLKGLPATEKSARYLTRGYSILIKDVTEEDAGNYTILLSIKQSNVFK 416
QY 411 NISLELVNVPQIHEKEASS---PSIYSHRSQALCTAYGVPPLSITQWHRPWTPCK 467
Db 417 NLATLVNVPKQIYKAVSFSDPDPALYPLGSRQILTAYGIPQ-P-TIKFWHPCNHNH 475
QY 468 MFAQSLRRQOQDLPOCRDRAVTTQDAVNPIESLDTWTEFVGEKNKTVSLVTONAN 527
Db 476 SEACDFCSNNEFSILD-----ADSNMGRIEISITQMAIEGKNKMASTLVVADSR 528
QY 528 VSAMKVCVSNKYGODERLIYFYVTIPDGF--TIESKPEELLEGPVLLSCQADSYKY 585
Db 529 ISGIYICIASNKVGTGVRNISFVITDPNGFHVNLKMP-T---EGEDLKSLCTVKNFLY 584
QY 586 EHLWYRL---NLSTLHDAGNPLLLDCKNVHLFATPLAASLEEVAPGARHA-TLSLSP 641
Db 585 RDVTWILLRTVNNNTMH-----YSISKOMATKEHSITLNTIM 624
QY 642 RVAPEHEGHVYCEVQDRRSHDKHCKKYLVSQALEAPLTONLTDLVNVSDSLEMOCIV 701
Db 625 NVSLQDSTYACRARNVTGTEILQKKEITRDQEAAPYLLRNLSDHTVAISSSTILDCHA 684
QY 702 AGAHAPSIVYKDERLLEKSGVDLADSNQKLSQRYREEDAGRYLCSVCNAKGVNSA 761
Db 685 NGVPEPQITWKNHKKIQOEPGIIILPGSSSTLFIERVTEDEGYVHCATNOKGSVESSA 744
QY 762 SVAVGSDKSGMEIVLGVGVIAVFEVFWLLLIIFCNRMPAHADIKTYLSIIMDPE 821
Db 745 YLTWGSTSKNLELITLCTCAATLFWLLTLIRKMKR-SSEIKTDYLSIIMDPE 803
QY 822 VPLEQCEYLSDASQWEPFRERHLGRVLYGAFGKVEASATGIHKGSCDVTAVKML 881
Db 804 VPLEDQERLPYDASKWNEFARERLKLKSLGRGAFGVVQASAFGIKKSPTCTRTAVKML 863
QY 882 KEGATASEHRLMSLKLTHIGHNLNVNLLGACTCPQGLMWIVFECYKGNLSNLA 941
Db 864 KEGATASEYKALMTELKLTTHIGHNLNVNLLGACTTQGGPLMWIVFECYKGNLSNLS 923
QY 942 KDAPF---SPCAEKSPEQGRFRAMVELARDRRRPGSSDRLVFARSKTEGGARRASPD 998
Db 924 KRDLFELNKAALHMEPKKMEPELEQKPRLDTSVTSSEFASGQEDKLSLSDVEE 983
QY 999 QEAEDLWLSPLTMEDLCYSFQVARGMEFLASRKCIIHDLAARNILLSSESQWVICDFGL 1058
Db 984 ESDGFGKEPITMEDLISYFQVARGMEFLASRKCIIHDLAARNILLSSENNVVICDFGL 1043
QY 1059 ARDIYKDPDVRKSGARLPLKWMAPESIFDKYITQSDVMSFVLLWEIFSLGASPPGV 1118

Db 1044 ARDIYKDPDVRKSGARLPLKWMAPESIFDKYITQSDVMSFVLLWEIFSLGASPPGV 1103
QY 1119 QINEEFQORLDRGTTRAPELATPAIRIMLNCWSDGDPKARPAFSELVEITGLDGLGRGL 1178
Db 1104 QMDDFCSRLREGMRRAPEYSTEIYQIMLDCWHDRPKRPRFAELVEKLGDLQANVO 1163
QY 1179 QEEVEVCMAPRSSOSSEGSFQSVSTMALHIAQDAED-----SPPSLQRHSLAARY 1231
Db 1164 QDQKD--YIPINAILTNGSFT-----YSTAFSEDFEKESISAPKFNSSGDDVRYV 1214
QY 1232 NWSFPCCLARGAETRGSSRMKTFEPPMPTTYKGSVDNQTSQGVNLAS---EPEQIE 1288
Db 1215 NAKNF-----MSLERIKTFEE--LLPNATSMFDDYQDSSITLASPMILKRFTWTD 1262
QY 1289 SRHR-----OESGFS-----CKPGONVAVTRAHPDSQGRRR 1320
Db 1263 SKPKASLKIDLRVTSKSGSLSDVSRPSCHSCGHV-----SEGRNR 1306
RESULT 10
160598
Fit-1 tyrosine kinase receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 24-Sep-1999
C:Accession: I60598
R:Yamane, A.; Seetharam, L.; Yamaguchi, S.; Gotoh, N.; Takahashi, T.; Neufeld, G.; Sh
Oncogene 9, 2683-2690, 1994
A:Title: A new communication system between hepatocytes and sinusoidal endothelial ce
A:Reference number: 160598; MUID:94336223
A:Accession: I60598
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1336 <RES>
A:Cross-references: GB:D28498; NID:g511662; PIDN:BAA03857.1; PID:g600379
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom
C:Keywords: ATP
F:825-1161/Domain: protein kinase homology <KIN>
F:833-841/Region: protein kinase ATP-binding motif
Query Match 33.9%; Score 2444.5; DB 2; Length 1336;
Best Local Similarity 40.5%; Pred. No. 8.3e-105;
Matches 572; Conservative 199; Mismatches 478; Indels 163; Gaps 37;
QY 5 AALCLRLWLCIGLDGLVSGYSMTPTLNTITEESHVIDTGDLSISCRGQHPLEAWPG 64
Db 8 AVLPCALLGL-LLTGYCSGSKLKGPELSLKGTHQINQAGOTLHOCRGAAHKS---62
QY 65 QEAPATGDKDSEDGVVYDCGTDPKCKVLLHVEHANDTGSVYCYIKARIEGT 124
Db 63 --LPTVISEDKLSVTSACGRNRRQFCSTLTLLNAQANHTGLYSCRYLPKSTSEKKM 120
QY 125 AASSVVFYRDFEQPFINK---PDTLLVNRKDMVWVCLVSLPGLNVTLS-OSSVLPW 179
Db 121 ESAIYIFSDAGSPFTEHSDIPKLVHTEGRELIIPCRVTSPTNITVLAKFPLDTLP 180
QY 180 QGEVVDORRGMVSTPLLDALYLOCTETWGDQDFLSNPLVHTGELNLYDIOLLPK 239
Db 181 GORIANDSRGFIISNATYKEIGLTCATVNGHLYKTN-YLTHROTNTIIDVOISPPSP 239
QY 240 LELLVGEKLVNCTVMAEFNSGVTPDWYPGKAERGKWPERRSQO-THEL-SSILPT 297
Db 240 VRELQGLTVNCTVTTDLNTRVQMSWNPYGRKATRAS-IRORIDQNSPHSNVFBVLKI 298
QY 298 HNVSHQDLGVSVCNANGIQRFESTEVIVHENFISVEMLKGPILFATAGDELKLPVK 357
Db 299 NNVESDKRGYTCVRSGSSTFTNTSVHVEKGFISVHRKQOQVETAGRSHLSWK 358
QY 358 LAAYPPPEFQYKDG---KALSGRHS-PH--ALVKEVTEASTGYTTLALMNSAAGLR 412
Db 359 VKAFSPPEVWLKDGVPATEKRSARYSVHGSYSLIKDVTEADAGDYTILLSIKQSLFRNL 418


```

QY 765 VEGSDKSGMEIVILVGTGVIATVFWFVLLLLIFCNMRRAHADIKTGVLSTIIMDPGEVPL 824
Db 746 VOGTSDKSNLELITITVTCVAATLFWLLTLIFIRKLK-SSSEVKTYDLSIIMDPDEVPL 804
QY 825 EQCELYSDASOWEPPRERLHGRVLYGAGFVQVAFSGIKHGKSSCDTVAVMKLEK 884
Db 805 DEQCELYSDASOWEPPRERLHGRVLYGAGFVQVAFSGIKHGKSSCDTVAVMKLEK 864
QY 885 ATASEHRALMSKLTILHGHNLVNVLLGACTKPPQPLVIVFECKYGNLSNLRKRD 944
Db 865 ATASEYKALMTKLTILHGHNLVNVLLGACTKPPQPLVIVFECKYGNLSNLRKRD 924
QY 945 AFSPCAEKSPE-----QRGFRAMVELARLDRRRPGSSDRVLFAF-----SKTEGAR 993
Db 925 LF--CLNKDAALHMLKESLEPLEGQOKPRLDSVSSSVTSFFEDRSVDSVEG--- 979
QY 994 RASPDQAEADLWLSPLTMEDLVYCSFQVARGMEFLASRKC IHRDLAARNILLSDDVYKI 1053
Db 980 ----DEDYSEISKQPLTMEDLVYCSFQVARGMEFLASRKC IHRDLAARNILLSDDVYKI 1035
QY 1054 CDFGLARDIYKDPYVKGSRARLPKMWAPESIFDKVYTTQSDVMSFGVLLWEIFSLGAS 1113
Db 1036 CDFGLARDIYKDPYVKGSRARLPKMWAPESIFDKVYTTQSDVMSFGVLLWEIFSLGAS 1095
QY 1114 PYPGVQINEEFCORLDRGTRMAPELATPAIRRMILNCWSDGPKARPAFSELVEITLGLL 1173
Db 1096 PYPGVQINEEFCORLDRGTRMAPELATPAIRRMILNCWSDGPKARPAFSELVEITLGLL 1155
QY 1174 OGRGLOEEVCMAPRSSQSSSEGSFQVSTMALHIAQDAEDSPSLQRHSLAARYNWM 1233
Db 1156 QANVOODGKD--YIPLNAILTRNSFTYSTPTFSDRGDKGFDAPHFHSGSSDDVRYVNA 1213
QY 1234 VSPGCLARGATRCSSRMKTFEFPMTPTTYKGSVDNQTDSGMVLAS--EEFEQIESR 1290
Db 1214 FRF-----MSLEKIKTFEE--LSPNSTSMFEDQIDTSTILGSLPRLRRFTWTETK 1261
QY 1291 -----HROESGFS 1298
Db 1262 PRASKMIDLRIASPRKEAGLS 1282

RESULT 12
B56598
endothelial kinase Quek2 - quail (fragment)
N:Alternate names: vascular endothelial growth factor receptor homolog Quek2
C:Species: Coturnix coturnix (quail)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 24-Sep-1999
C:Accession: B56598
R:Eichmann, A.; Marcelle, C.; Breant, C.; Le Douarin, N.M.
Mech. Dev. 42, 33-48, 1993
A:Title: Two molecules related to the VEGF receptor are expressed in early endothelial
A:Reference number: A56598; MUID:93378866
A:Contents: ssp. japonica, E4 embryo
A:Accession: B56598
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-367 <EID>
A:Cross-references: GB:S65207; NID:9410682; PIDN:AAB28128.1; PID:g410683
A:Note: sequence extracted from NCBI backbone (NCBIN:137164, NCBI:P:137165)
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; growth factor receptor
F:1-176/Domain: protein kinase homology (fragment) <KIN>

Query Match 17.8%; Score 1283.5; DB 2; Length 367;
Best Local Similarity 69.8%; Pred. No. 3.2e-52;
Matches 256; Conservative 32; Mismatches 66; Indels 13; Gaps 5;

QY 1003 DLWLSPLTMEDLVYCSFQVARGMEFLASRKC IHRDLAARNILLSDDVYKICDFGLARDI 1062
Db 3 DLWQSPLTMLDLYCSFQVARGMEFLASRKC IHRDLAARNILLSDDVYKICDFGLARDI 62
QY 1063 YKDPDYVKGSRARLPKMWAPESIFDKVYTTQSDVMSFGVLLWEIFSLGASYPGPVQINE 1122

```

```

Db 63 YKDPDYVKGSRARLPKMWAPESIFDKVYTTQSDVMSGLVLLWEIFSLGASYPGPVQINE 122
QY 1123 EFCORLDRGTRMAPELATPAIRRMILNCWSDGPKARPAFSELVEITLGLDQGRLOEEZ 1182
Db 123 EFCORLDRGTRMAPELATPAIRRMILNCWSDGPKARPAFSELVEITLGLDQGRLOEEZ 181
QY 1183 EVCMAFRSSOSSEGSFQVSTMALHIAQDAEDSPSLQRHSLAARYNWMVSPGCLAR 1242
Db 182 KDYIPLNDSSHSSDDGFSQVPSAAQ--NSDEEDFMRIRCHSLAARYNCVSPGCLTG 239
QY 1243 GASTRGSSRMKTFEFPMTPTTYKGSVDNQTDSGMVLASEEFPQIESRHRQESGFSCKGP 1302
Db 240 GNOIRCSSRIKTFEFPMTHTYKAHPDNOTDSGMVLASEEFPQIESRHRQESGFSCKGP 299
QY 1303 QONVAVTRAHPDSQRRRPERGAR--GGVTFYNSVEGSELSEPEDEHCHSPSA----- 1353
Db 300 NRTAELSAEQSDLRG--RCRPSYQVGGQTFYNSVEGSELSEPEDEHCHSPSA----- 358
QY 1354 RVTFEFD 1360
Db 359 HASFFSE 365

RESULT 13
PFHUGA
platelet-derived growth factor receptor alpha precursor - human
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 11-Jun-1999
C:Accession: A40162; A32941
R:Matsui, T.; Heldaran, M.; Miki, T.; Popescu, N.; La Rochelle, W.; Kraus, M.; Pierce
Science 243, 800-804, 1989
A:Title: Isolation of a novel receptor cDNA establishes the existence of two PDGF rec
A:Reference number: A40162; MUID:89130149
A:Accession: A40162
A:Molecule type: mRNA
A:Residues: 1-1089 <MATS>
A:Cross-references: GB:M21574; NID:g189733; PIDN:AAA96715.1; PID:g189734
R:Claesson-Welsh, L.; Eriksson, A.; Westermarck, B.; Heldin, C.H.
Proc. Natl. Acad. Sci. U.S.A. 86, 4917-4921, 1989
A:Title: cDNA cloning and expression of the human A-type platelet-derived growth fact
A:Reference number: A32941; MUID:89296915
A:Accession: A32941
A:Molecule type: mRNA
A:Residues: 1-1089 <CLA>
A:Cross-references: GB:M22734; NID:g189725; PIDN:AAA60048.1; PID:g189726
C:Comment: The extracellular domain is predicted to include five immunoglobulin-like
C:Genetics:
A:Gene: GDB:PDGFRA
A:Cross-references: GDB:I20267; OMIM:173490
A:Map position: 4q11-q12
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo
C:Keywords: ATP; autophosphorylation; dimer; glycoprotein; growth factor receptor; ph
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-1089/Product: platelet-derived growth factor receptor alpha #status predicted <EXT>
F:25-524/Domain: extracellular #status predicted <EXT>
F:42-103/Domain: immunoglobulin homology <IMM1>
F:143-191/Domain: immunoglobulin homology <IMM2>
F:238-292/Domain: immunoglobulin homology <IMM3>
F:428-503/Domain: immunoglobulin homology <IMM4>
F:525-548/Domain: transmembrane #status predicted <TMH>
F:549-1089/Domain: intracellular #status predicted <INT>
F:591-957/Domain: protein kinase homology <KIN>
F:599-607/Region: protein kinase ATP-binding motif
F:42,76,103,179,353,359,458,468/Binding site: carbohydrate (Asn) (covalent) #status p
F:49,100,150,189,235,290,435-501/Disulfide bonds: #status predicted
F:627/Active site: Lys #status predicted
F:849/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi

Query Match 16.1%; Score 1163.5; DB 1; Length 1089;
Best Local Similarity 26.6%; Pred. No. 3.3e-46;

```


Tue Jan 23 09:28:37 2001

F:848/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 16.0%; Score 1151; DB 1; Length 1088;
Best Local Similarity 26.2%; Pred. No. 1.2e-45;
Matches 360; Conservative 184; Mismatches 395; Indels 434; Gaps 46;

QY 17 LLDGLVSGVMP-----PTLNITEESHVIDTGDLSISCRGHPLEWAWPQAEATG 71
DB 9 VLSCLLTGSLVLCOLLPSILPNEKIVPLSSSFLRCFGESEVSWHPMSEE----- 63
QY 72 DKDSGDTGVVRCDEGTDARPYCKVLLHEVHANDTGVVYKYIKARIEGTTAASSVVF 131
DB 64 -----EDPNEVEIRSENNSGLVTVLEVNNASAAHTGWTTCYNNHTQTESEIEGRHIY 119
QY 132 VRDFEQPFI--NKPDTL-LVNRKDAWVPCVLSIPGLNVTLSQSLSVLPDQGVVWDDR 188
DB 120 VPDPMFVPLGMDSLVIVEEDSAIIPCLTDPDTEVTLHNNGLV-----PASYSR 174
QY 189 RGLMVLSTPLLDALYLQCTTGGDDFLSNPFLVHTGNELVDIOLLPKLSLEL----- 242
DB 175 QGF---NGTFSVGPYI-CEAAVKGRFTKSEF-----NVYALKATSELNLEMDROTQV 223
QY 243 -LVGEKLVNCTVWAEFNSGVTFDMVDPKQAEKRWKPERKSQOHTHELSSI-----LT 296
DB 224 YKAGETIVTCAVEN--NEVVDLQWTPGVRNKGITMLE-----EIKLPSIKLVITLT 275
QY 297 IHNVSQHDLSYVCKANNGIQPRE--STEVIVHENPFISVEMWLGPILEATAGDELVKL 354
DB 276 VPKATVNDSDGYEACAAQATKVEKMTVTISVHEKGFVQIRPTTGH-LETNVLHQVREF 334
QY 355 PVKLAAYPPPEFYQWYDKGKALSGRSHALVLEKVT-----LTEITDVORSGOETRYQSKLIRAKED 387
DB 394 TGYITLWNSAAGLRNLSLELVNVPQIHEKEASSPISYRHSR-----QALTATAGV 450
DB 388 SGHYITIVQDDP--MKSYTFELSTLVPASILE-----LVDDHGGSGGQVTRCTAET 439
QY 451 PLPLSIQHWRRPWPCKMFAQRSLRRRQOQDLMPQCRD---WRAVTTODAVNPFIESLDTW 507
DB 440 PLP-NIE-----WIKCD-----IKKCNNTSM-TVLASVNSNII----- 472
QY 508 TEFEV-GKNTVSKLVTONANVAMKYCVSVNKGQDERLIFYVTTPIDGFTIESKPSE 566
DB 473 TEFHQGRSTVEGRSVPAKVEETIAVRCLAKNDLGIGNR----- 511
QY 567 ELLEGQPVLLSCQADSVKYEHLRWYRLNLTSLDHAHGNPLLLDCKNVLHLPATPLAASLEE 626
DB 512 -----ELKL 515
QY 627 VAPGARHATLSLTPRAPEHEGHVCEVQDRSHDKHKHYLSVQALEAPRLTONLTD 686
DB 516 VAPSLR-----SELTAAAVLVLLVIVIVS 540
QY 687 LLVNVSDSLEMQCLVAGAHAPSVIWKDERLLEKSGVDLADSNOKLSIORVREEDAGRY 746
DB 541 LIVLV-----VLWK-----OKPYEIVRWV 560
QY 747 LCSVCNAKGCYNSSAVAGESEDKGSNEIIVLVGTGVIAVFFWVLLLLIFCNRRAHA 806
DB 561 IESI-----SPD----- 567
QY 807 DIKGTLSIIMDPCVEPLEEOCEVLSYDASQWEPFRERHLGRVLYGAFGKVEASAFG 866
DB 568 -----GHEYIYVDPQLP-----YD-SWEEFPRDLVGLRILGSGAFGKVEGTAYG 613
QY 867 IHKGSCTDVAVKMLKEGATASEHRLMSLEKLIHIGNHLNVNVLNVLGACTKPGQPLMWI 926
DB 614 LRSRQPVKVAVKMLKPTARSSEKQALMSLEKIMTHGLPHNLNVLNVLGACTK-SGPIYII 672
QY 927 VEFCKYGNLSNFLRAKRDAP-----SPCAEK-----SPEORGFRFA 962
DB 927 VEFCKYGNLSNFLRAKRDAP-----SPCAEK-----SPEORGFRFA 962

DB 673 TEYCFYGDVLNVHLNKRDSFMSRHHPEKPKKDLDFGLNPADESTRSYVILSPENNGDYVD 732
QY 963 WVE-----LALDRRPGSSDRVLFARFSKTEGGARRASPDQAEEDWLSP-----LTM 1011
DB 733 MKQADTTQVPLMLERKEVSKYSDIQKSLYDRPASTYKKKSHLSEAKNL-LSDDSEGLTL 791
QY 1012 EDLVCSYFQVARGMEFLASRKICHRDLAARNILLSESDVYKICDFGLARDIYKDPYRK 1071
DB 792 LDLLSTYQVARGMEFLASKNCVHRDLAARNVLLAQKIVKICDFGLARDIMHDSNYVSK 851
QY 1072 GSARLPLKMAPESIDFKVYTTQSDVMSGVLLWEIFSLGASPYGCVQVINESFCORLDD 1131
DB 852 GSTFLPVKMAPESIDFNLYTLLSDVMSGVLLWEIFSLGASPYGCVQVINESFCORLDD 911
QY 1132 TRARAPETAPARRITMLNCWGDPKARPAFSELVILGDLQGRLOEVEEVCNAPRESS 1191
DB 912 YRAKPDHATSEVVEIMVQCWNEPEKRPFSYHLSLSEVENLLPGQYKKSVEKHL----- 966
QY 1192 OSSEEGFSQVSTMAHIAQDAEDSPPSLQHSLAARYNWNVSFFPGCLARGAETRGSSR 1251
DB 967 -----DFLASDHPAVARMRVSDNA-----YIGVTKN-----EEDK 998
QY 1252 MKTEFEFPMTPTYKGSVDNQ---TDSGMVL-----ASEFEQIESRHRQES 1295
DB 999 LKWE-----GGLEQRLSADSGYIIPLPDIDPVEEEDLGRNRHSS 1041

RESULT 15
S33727
platelet-derived growth factor receptor alpha precursor ' mouse
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Mus musculus (house mouse)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Mar-2000
C:Accession: I57511; S33727
R:Stiles, C.D.; Wang, C.
Mol. Cell. Biol. 10, 6781-6784, 1990
A:Title: Retinoic acid promotes transcription of the platelet-derived growth factor a
A:Reference number: I57511; MUID:91061789
A:Accession: I57511
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Peptides: 1-1089 <RES>
A:Cross-references: GB:M57683; NID:g199783; PIDN:AAA39733.1; PID:g199784
C:Genetics:
A:Gene: PDGF-alpha-R
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo
C:Keywords: ATP; glycoprotein; growth factor receptor; phosphotransferase; transmembr
F:1-23/Domain: signal sequence #status predicted <SIG>
F:228-292/Domain: immunoglobulin homology <IMH>
F:591-957/Domain: protein kinase homology <KIN>
F:599-607/Region: protein kinase ATP-binding motif
F:42,76,89,103,179,353,359,458,468,506/Binding site: carbohydrate (Asn) (covalent) #s

Query Match 15.6%; Score 1125; DB 1; Length 1089;
Best Local Similarity 26.9%; Pred. No. 1.9e-44;
Matches 369; Conservative 178; Mismatches 394; Indels 430; Gaps 51;

QY 14 CIGLLDGLVSGVSTPPTLNITEESHVIDTGDLSISCRGHPLEWAWPQAEATGDK 73
DB 13 CLTQFGLISCOLLPSILP-NEKEIVQLNSSFSLRCVGESEVSWHPMSEE----- 64
QY 74 DSEDGTGVVRCDEGTDARPYCKVLLHEVHANDTGVVYKYIKARIEGTTAASSV 130
DB 65 --EDPNEVEIRSENNSGLVTVLEVNNASAAHTGWTTCYNNHTQTESEIEG---RHIYI 119
QY 131 FYRDEQFPI--NKPDTL-LVNRKDAWVPCVLSIPGLNVTLSQSLSVLPDQGVVWDD 187
DB 120 YPPDPMFVPLGMDSLVIVEEDSAIIPCLTDPDTEVTLHNNGLV-----PASTDS 174
QY 188 RGLMVLSTPLLDALYLQCTTGGDDFLSNPFLVHTGNELVDIOLLPKLSLEL----- 242
DB 175 ROGF---NGTFSVGPYI-CEAAVKGRFTKSAF-----NVYALKATSELNLEMDARQT 223

QY 1254 TFEFPMTPTTYKGSVDNQ---TDSGMVL-----ASEEPEQIESREROES 1295
Db 1002 DWE-----GGLEQRLSADSGYIIPDPIDPVPPEEDLGKRRHSS 1042

Search completed: January 23, 2001, 05:04:36
Job time: 164 sec

QY 243 --LVGEKLVNCTVWAEFNSGVTFDMDYPGKQAEGRKVPERRSQOHTHELSSI-----L 295
Db 224 VTKAGETIVTCAVFN--NEVVDQWTFPGGVRNGITMLE-----EIKLPSKVVTYL 275
QY 296 TIHNSQHDLSYVCKANQIREFRESTEVI--VHENFFISVEMWLGPI---LEATAGDE 350
Db 276 TYPKATVDSGEYECQAARQATKEVKEMKRVITISVHEKGEVEIE---PTFSQLEPVNLHE 331
QY 351 LVKLPVKLAAYPPPEFYQKDG-----KALSGRHSFHALVLKVEYTEASTG 395
Db 332 VREFVVEQAYPTPRISWLNKONLTIENLTIITDQVKSQETRYQ-SKLKLIRAKEEDSG 390
QY 396 TYTLALMSAAGLRNRNLSLELVNVPQIHEKEASSPSIYSRH---SROALTCTAYGVP 451
Db 391 HTYIIQNEDD--VASYTFELSTLP-----ASILDVDDHGGGGQIVRCTAEGP 441
QY 452 LPLSIOMHWRPWTCKMFAQRSLRRRQOQDLMPQCRDHRVTTODAVNPIESLDT-WTEF 510
Db 442 LPEI-----DMWICKH-----IKKCN-----DTSMT-- 463
QY 511 VEGKNTVSKLVIONANYSAMKYCVSNKVGODERLIYFYVTTIP--DGFTIESKPSEEL 568
Db 464 -----VLASNSVNI-----ITELPRGRSTVEGRVSFAK 492
QY 569 LEGQPVLLSCQADSYKYEHLRWYRLNLSTLHDAGNPLLLDCKKNVHLFATPLAASLEEVA 628
Db 493 VE-ETIAVRCLAKN-----NLSV-----VARELKIVA 518
QY 629 PGARHATLSLTPRVAPEHEGHYCEVQDRSHDKHKKYLSVOALEAPRLTONLTDLL 688
Db 519 PTLR-----SELTVAEAVLVLLVIVISLI 543
QY 689 VNVSDSLEMOCLVAGAHAPSIVWYKDERLLEKSGVDLADSNOKLSIORVREEDAGRYLC 748
Db 544 VLV-----VWK-----OKPRYEIRWVIE 563
QY 749 SVCNAGCVNNSASVAVGSEDKSGMEIVILVGTGVIAVFTFWLLLLIFCNRRPAHADI 808
Db 564 SI-----SPD----- 568
QY 809 KTYLSIIMDGPVPLEQCEVLSVDASOMEPFRERLHLGRVLGYGAEKVVESAFGIIH 868
Db 569 --GHEYIYDPMQLP-----YD-SRWEPRDGLVGRILGSGAFGKVVGTAYGLS 616
QY 869 KGSCTVAVKMLKEGATASEHRALMSELKILIHIGNHLNVNVLGACTKPGQPLMWIVE 928
Db 617 RSOPVMKVAVKMLKPTARSSEKQALMSELKINTHLGPHNLVNLGACTK-SGPIIITE 675
QY 929 FCYGNLSNFLRAKRDAP-----SPCAEK-----SPEQRGRFRAMV 964
Db 676 YCFYGLVNLHKNRDSFMSQHPKPKDLOIFGLNPADESTRSYVILSFENNGDYMDMK 735
QY 965 E-----LARDRRRPGSDRVLFARFSKTEGGARRASPDQAEIDLWLSL-----LTWED 1013
Db 736 QDDTQYVPMLEKKEVSKYSIORSLYDRPASYKKKMLDSEVKNL-LSDDDSGLTLTD 794
QY 1014 LVCYSQVARGMEFLASRCKIHRDLAARNILSESDVYKICDFGLARDIYKDPDYVRKGS 1073
Db 795 LLSFTYQVARGMEFLASKNCHVRDLAARNVLLAQCKIVKICDFGLARDIMHDSNDVSKGS 854
QY 1074 ARPLKMAPEISIFDKVTTQSDVNSFGVLLWEIFSLGASVPYGVQINEEFQORLDCGR 1133
Db 855 TELPKWMAPEISIFNLNLTSLDWSYGLLWEIFSLGTPYPGMNVDSFTYKIKSGYR 914
QY 1134 MRAPELATPAIRIMLNCWSDPKRPAFSELVEILGDLQGRGLQEEVEYCMAPRSSQS 1193
Db 915 MAKPOHATSEVYEIMVQCNWNSDPKRPSPFYHLSEILENLLPGQYKKSYEKIHL----- 967
QY 1194 SEEGSFQSVSTWALHIAQADAEDSPSLQRHSLAARYNWTVSFPCGLARGAETRGSSRMK 1253
Db 968 ----DFLKSDBPAVARMRVDSNA-----YIGVTYKN-----EEDKLK 1001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 23, 2001, 05:03:27 ; Search time 65 Seconds
(without alignments)
677.182 Million cell updates/sec

Title: US-09-375-248-2
Perfect score: 7211
Sequence: 1 MORGALCLRLWLCGLLDG.....SEEDHCSPSARVTFEFTDINSY 1363

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues
Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	6852	95.0	VGR3_HUMAN	P35916 homo sapien
2	6320	87.6	VGR3_MOUSE	P35917 mus musculus
3	2804	38.9	VGR2_COTJA	P35983 coturnix co
4	2762	38.3	VGR2_HUMAN	P35968 homo sapien
5	2669	37.0	VGR2_MOUSE	P35918 mus musculus
6	2646	36.7	VGR2_RAT	O08775 rattus norv
7	2469	34.2	VGR1_MOUSE	P35969 mus musculus
8	2463	34.2	VGR1_HUMAN	P17948 homo sapien
9	2444.5	33.9	VGR1_RAT	P53767 rattus norv
10	1163.5	16.1	PGDS_HUMAN	P16234 homo sapien
11	1160	16.1	PGDS_MOUSE	P26618-mus musculus
12	1151	16.0	PGDS_RAT	P20786 rattus norv
13	1122.5	15.6	PGDS_XENLA	P26619 xenopus lae
14	1084	15.0	PGDR_HUMAN	P09619 homo sapien
15	1052.5	14.6	PGDR_MOUSE	P05622 mus musculus
16	1042.5	14.5	KKIT_HUMAN	P10721 homo sapien
17	1033.5	14.3	KKIT_MOUSE	P05532 mus musculus
18	1027.5	14.2	KKIT_CAPHI	Q28317 capra hircu
19	1024	14.2	KKIT_BOVIN	P43481 bos taurus
20	1021.5	14.2	KKIT_CHICK	Q08156 gallus gall
21	1014	14.1	KKIT_FELCA	Q28889 felis silve
22	997	13.8	KKIT_CANFA	O97799 canis fami
23	967	13.4	KKFS_FELCA	P13369 felis silve
24	956	13.3	KFMS_FSVMD	P00545 feline sarc
25	954.5	13.2	KFMS_HUMAN	P07333 homo sapien
26	954.5	13.2	FLT3_MOUSE	Q00342 mus muscu
27	951.5	13.2	KFMS_MOUSE	P09581 mus muscu
28	949	13.2	KFMS_RAT	Q00495 rattus norv
29	904.5	12.5	FLT3_HUMAN	P36888 homo sapien
30	896	12.4	CEK2_CHICK	P18460 gallus gall
31	869	12.1	FGR1_HUMAN	P11362 homo sapien
32	867.5	12.0	KKIT_FSVH2	P04048 feline sarc
33	867	12.0	FGR1_MOUSE	P16092 mus musculus

34 862 12.0 822 1 FGR1_RAT
35 858.5 11.9 806 1 FGR3_HUMAN
36 854.5 11.8 819 1 FGR1_CHICK
37 852 11.8 812 1 FGR1_XENLA
38 844 11.7 801 1 FGR3_MOUSE
39 842 11.7 821 1 FGR2_HUMAN
40 840.5 11.7 654 1 BFR2_MOUSE
41 835 11.6 821 1 FGR2_MOUSE
42 833.5 11.6 823 1 CEK3_CHICK
43 830.5 11.5 813 1 FGR2_XENLA
44 824.5 11.4 802 1 FGR4_HUMAN
45 812.5 11.3 808 1 FGR4_MOUSE

O04589 rattus norv
P22607 homo sapien
P21804 gallus gall
P22182 xenopus lae
Q01851 mus musculus
P21802 homo sapien
Q01742 homo sapien
P21803 mus musculus
P18461 gallus gall
Q03384 xenopus lae
P22455 homo sapien
Q03142 mus musculus

ALIGNMENTS

RESULT 1
VGR3_HUMAN
ID VGR3_HUMAN STANDARD; PRT; 1298 AA.
AC P35916;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 3 PRECURSOR (EC 2.7.1.112)
DE (VEGFR-3) (TYROSINE-PROTEIN KINASE RECEPTOR FLT4).
GN FLT4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=93241723; PubMed=8386825;
RA Galland F., Karanysheva A., Pebusque M.-J., Borg J.-P., Rottapel R.,
RA Dubreuil P., Rosnet O., Birnbaum D.;
RT "The FLT4 gene encodes a transmembrane tyrosine kinase related to the
RT vascular endothelial growth factor receptor.";
* RL Oncogene 8:1233-1240(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93007956; PubMed=1327515;
RA Pajusola K., Aprelikova O., Korhonen J., Kaipainen A., Pertovaara L.,
RA Alitalo R., Alitalo K.;
*FLT4 receptor tyrosine kinase contains seven immunoglobulin-like
loops and is expressed in multiple human tissues and cell lines.";
Cancer Res. 52:5738-5743(1992).
RN [3]
RP ERRATUM.
RA Pajusola K., Aprelikova O., Korhonen J., Kaipainen A., Pertovaara L.,
RA Alitalo R., Alitalo K.;
Cancer Res. 53:3845-3845(1993).
RN [4]
RP SEQUENCE FROM N.A.
RA Lee J., Gray A., Yuan J., Luoh S.M., Avraham H., Wood W.I.;
Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 761-1190 FROM N.A.
RX MEDLINE=92119639; PubMed=1310071;
RA Aprelikova O., Pajusola K., Partanen J., Armstrong E., Alitalo R.,
RA Bailey S.K., McMahon J., Wasmuth J., Huebner K., Alitalo K.;
*FLT4, a novel class III receptor tyrosine kinase in chromosome 5q33-
qter.";
Cancer Res. 52:746-748(1992).
CC -!- FUNCTION: RECEPTOR FOR VEGF-C. HAS A TYROSINE-PROTEIN KINASE
CC ACTIVITY.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: PLACENTA, LUNG, HEART, AND KIDNEY, DOES NOT
CC SEEM TO BE EXPRESSED IN PANCREAS AND BRAIN.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement ([see http://www.isb-sib.ch/announcement/](http://www.isb-sib.ch/announcement/) or send an email to license@sib-sib.ch).

[illegible]

Query Match

95.0%; Score 6852; DB 1; Length 1298;

Db 1021 VARGMEFLASRKCIRHDLAARNILLSSESDVVKICDGLARDIYKDPDYVRKGSARLPLK 1080
 Qy 1081 MAPESIFDKVYTTSDVWSFGVLLWEIFSLGASYPGVQINEEFCQRLDGTMRAPELA 1140
 Db 1081 MAPESIFDKVYTTSDVWSFGVLLWEIFSLGASYPGVQINEEFCQRLDGTMRAPELA 1140
 Qy 1141 TPAIRIMLNCWSGDKPRAPAFSELVELLDLGRGLQEEVEVCMAPRSSQSEGSFS 1200
 Db 1141 TPAIRIMLNCWSGDKPRAPAFSELVELLDLGRGLQEEVEVCMAPRSSQSEGSFS 1200
 Qy 1201 QVSTMALHIAQDAEDSPSLQRSLAARYNWNYSFPGCLARGAETRGSRMKTFEFP 1260
 Db 1201 QVSTMALHIAQDAEDSPSLQRSLAARYNWNYSFPGCLARGAETRGSRMKTFEFP 1260
 Qy 1261 TPTTYKGSVDNOTSGMVLASEEPEQIESRHOESGF 1297
 Db 1261 TPTTYKGSVDNOTSGMVLASEEPEQIESRHOESGF 1297

RESULT 2

VGR3_MOUSE STANDARD; PRT: 1363 AA.

AC P35917;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 3 PRECURSOR (EC 2.7.1.112)
 DE (VEGFR-3) (TYROSINE-PROTEIN KINASE RECEPTOR FLT4).
 GN FLT4 OR FLT-4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RC STRAIN=C57BL/6J;
 RA MEDLINE=9330572; PubMed=8393164;
 RA Finnerty H., Kelleher K., Morris G.E., Bean K., Merberg D.M.,
 RA Kriz R., Morris J.C., Sookdeo H., Turner K.J., Wood C.R.;
 RT "Molecular cloning of murine FLT and FLT4.";
 RL Oncogene 8:2293-2298(1993).

CC -!- FUNCTION: RECEPTOR FOR VEGF-C. HAS A TYROSINE-PROTEIN KINASE

CC ACTIVITY.

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT LUNG AND LIVER, AND IN

CC FETAL LIVER, BRAIN, INTESTINE AND PLACENTA.

CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.

CC -!- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: L07296; AAA00077.1; -

DR HSSP; P11362; 1FGI.

DR MGD; MGI:95561; FLT4.

DR INTERPRO: IPR000719; -

DR INTERPRO: IPR001245; -

DR INTERPRO: IPR001824; -

DR INTERPRO: IPR003006; -

DR PFAM; PF00047; Ig; 6.

DR PFAM; PF00069; pkinase; 2.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

KW Transferase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;

KW Receptor; Transmembrane; Signal; Immunoglobulin domain; Glycoprotein.

FT SIGNAL 1 24

POTENTIAL.

FT CHAIN 25 1363 VASCULAR ENDOTHELIAL GROWTH FACTOR
 FT RECEPTOR 3.
 FT DOMAIN 25 775 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 776 797 POTENTIAL.
 FT DOMAIN 798 1363 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 44 118 IG-LIKE DOMAIN 1.
 FT DOMAIN 151 213 IG-LIKE DOMAIN 2.
 FT DOMAIN 245 317 IG-LIKE DOMAIN 3.
 FT DOMAIN 351 403 IG-LIKE DOMAIN 4.
 FT DOMAIN 438 541 IG-LIKE DOMAIN 5.
 FT DOMAIN 571 660 IG-LIKE DOMAIN 6.
 FT DOMAIN 692 758 IG-LIKE DOMAIN 7.
 FT DOMAIN 845 1173 PROTEIN KINASE.
 FT NP_BIND 851 859 ATP (BY SIMILARITY).
 FT BINDING 879 879 ATP (BY SIMILARITY).
 FT ACT_SITE 1037 1037 BY SIMILARITY.
 FT DISULFID 51 111 POTENTIAL.
 FT DISULFID 158 206 POTENTIAL.
 FT DISULFID 252 310 POTENTIAL.
 FT DISULFID 445 534 POTENTIAL.
 FT DISULFID 578 653 POTENTIAL.
 FT DISULFID 699 751 POTENTIAL.
 FT MOD_RES 1068 1068 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 33 33 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 104 104 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 166 166 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 251 251 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 299 299 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 411 411 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 527 527 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 582 582 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 594 594 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 683 683 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 690 690 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 758 758 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 1363 AA; 153015 MW; F1BF8A2BDEF99BE9 CRC64;

Query Match 87.6%; Score 6320; DB 1; Length 1363;

Best Local Similarity 87.5%; Pred. No. 0;

Matches 1192; Conservative 58; Mismatches 103; Indels 0; Gaps 0;

Qy 1 MORGAAALRLWLCLGLLDGLVSGYSMTPTLNITEESHVIDTGDLSISCRGQPLEWA 60

Db 1 MOPGAALNLWLCLGLGLQGLANGYSMTPTLNITEESYVIDTGDLSISCRGQPLEWT 60

Qy 61 WPGAQAPATGDKRDS EDTGVVROCEGTDARPCKVILLHEVHANDTGSYCYKIKARI 120

Db 61 WPGAQEVLTITGGKDS EDTRVVHDCGTEARPYCKVLLAOTHAANTGSYHCYKIKARI 120

Qy 121 EGTAAASYVVRDFEQPFINKPDTLLVNRKRDAMVPCVLSIPGLNLTLSQSQSVLPD 180

Db 121 EGTAAASYVVRDFEQPFINKPDTLLVNRKDSHWVPCVLSIPGLNLTLSQSQSVLPD 180

Qy 181 QEVVMDRRRGMVSTPLLDHALYLQCTETGDDQDFLSNPFLVHTITGELDYDIQLPK 240

Db 181 QEVLDVDRRGMVSTPLLDHALYLQCTETGDDQDFLSNPFLVHTITGELDYDIQLPK 240

Qy 241 ELLVGEKVLNCTVWAEFNSGVTFDWDY PKQAEGRKLVWVRRSQQTETELSSILT 300

Db 241 ELLVGEKVLNCTVWAEFNSGVTFDWDY PKQAEGRKLVWVRRSQQTETELSSILT 300

Qy 301 SQHDLGSYVCKANNIGIQRRESTEVIYVHEPFIISVEVLKGFILAEATAGDELKLPV 360

Db 301 SQHDLGSYVCKANNIGIQRRESTEVIYVHEPFIISVEVLKGFILAEATAGDELKLPV 360

Qy 361 YPPEFQWYKDGKALSGRHSHPHALVLEKVEASTGTVTLWNSAAGLRNLSILELV 420

Db 361 YPPEFQWYKDKKAVTGRHNHPHALVLEKVEASTGTVTLWNSAAGLRNLSILELV 420

Qy 421 PPOIHEKAEASSPSYSRHSRQALCTCTAYGVPLPLSIQHWHRPWTCKMFAQSR 480

Db 421 PPOIHEKAEASSPSYSRHSRQALCTCTAYGVPLPLSIQHWHRPWTCKMFAQSR 480

Db 421 PPHIHEKEASSPSIYSHRSQTLLCTAYGVQPLSVQWHRPWPCKTFAQSLRRRQOR 480
 Qy 481 DLMPQCRDRAVTTQDAVNIIESLDITTEFEVGEKNKTVSKLVTONANVSAMYKCVYNKV 540
 Db 481 DGMFQCRDREVTQDAVNIIESLDSWTEFEVGEKNKTVSKLVQDANVSAMYKCVYNKV 540
 Qy 541 GODERLIYFVYTTIPDGFTEKSPSELEGGOPVLLSCQADSKYKHLRWYRLNLSLTHD 600
 Db 541 GODERLIYFVYTTIPDGFTEKSPSELEGGOPVLLSCQADSKYKHLRWYRLNLSLTHD 600
 Qy 601 AHGNPLLDCKNHLFATPLAASLEEVAPGARHATLSLSPRAPEHEGHYVCEVQDRRS 660
 Db 601 AOGNPLLDCKNHLFATPLAANLEEAEPGARHATLSLSPRAPEDEGDYVCEVQDRRS 660
 Qy 661 HDRCHKKYLSVQALEAPRLTQNTDLLVNVSDSLEQCLVAGAHAPSIVWYKDERLEE 720
 Db 661 QDRCHKKYLSVQALEAPRLTQNTDLLVNVSDSLEMRCPVAGAHAPSIVWYKDERLEE 720
 Qy 721 KSGVDLADSNOKLSIQVRREDAGRYLCSVCNAKGCYNSSASVAVESGDKSMEIVILI 780
 Db 721 ESGIDLADSNORLSIQVRREDAGRYLCSVCNAKGCYNSSASVAVESGDKSMEIVILI 780
 Qy 781 GTGVIAVEFWLLLIIFCNMRRAHADIKTGYLSIIMDPGEVPLEEOCYLSYDASOWEF 840
 Db 781 GTGVIAVEFWLLLIIFCNMRRAHADIKTGYLSIIMDPGEVPLEEOCYLSYDASOWEF 840
 Qy 841 PRERHLGRVLGYGAFGKVVPEASAFGTHKGSSCDTAVKMLKEGATASEHRALMSLKIL 900
 Db 841 PRERHLGRVLGYGAFGKVVPEASAFGTHKGSSCDTAVKMLKEGATASEHRALMSLKIL 900
 Qy 901 IHIGNHLNVNLLGACTKPGPLMIVIEFCYKGNLSFLRAKRDGAFSPCAEKSPEQRGRF 960
 Db 901 IHIGNHLNVNLLGACTKPGPLMIVIEFCYKGNLSFLRVKRDTFNPYAEKSPQRRRF 960
 Qy 961 RAMVELARLDRRRGSSDRVLFAFESKTEGGARRASPDQEAEDLWLSPLTHEDLVCSYFQ 1020
 Db 961 RAMVEGAKADRRRGSSDRALFTRFLMGKSARRAPLVQEAEDLWLSPLTHEDLVCSYFQ 1020
 Qy 1021 VARGMEFLASRCKIHRDLAARNILLSDDVYKICDFGLARDIYKDPYVRKGSARLPKW 1080
 Db 1021 VARGMEFLASRCKIHRDLAARNILLSDDVYKICDFGLARDIYKDPYVRKGSARLPKW 1080
 Qy 1081 MAPESIFDKVYTTQSDVMSFVGLWEIFSLGASYPGVQINEEFCQRLKDGTRAPELA 1140
 Db 1081 MAPESIFDKVYTTQSDVMSFVGLWEIFSLGASYPGVQINEEFCQRLKDGTRAPELA 1140
 Qy 1141 TPAIRHMLNCSGDPKARPAFSELVEILGDLQGRGLQEEVEECMAPRSSQSSEGSFS 1200
 Db 1141 TPAIRHMLNCSGDPKARPAFSDULVEILGDLQGGQWQEEEREMALHSSQSSEEDGFM 1200
 Qy 1201 QYSTMALHTAADAEDSPSLQRHSLAARYNWNVSFPGLARGAETRGSSRMKTFEEFPM 1260
 Db 1201 QASTTALHTEADADDSPSMHCHSLAARYNCVSPGLARGTGTGSSRMKTFEEFPM 1260
 Qy 1261 TPTTYKGSVDNQTDSGMYLASEEFPQIESRHQESGFSCKGPGQNVAVTRAHPDSQGRRR 1320
 Db 1261 TPTTYKASNDNQTDSGMYLASEEFPQIESRHRPEGSFCKGPGQNDIPRGHPDPQGRRR 1320
 Qy 1321 RPERGARGOVFNSEYGELESPSEDEHCSPSARVYTFDTSY 1363
 Db 1321 RPTQAGGKGVYNNYGEVQPCTEGDCPSAGSTFFADSSY 1363

RESULT 3
 VGR2_COTJA
 ID VGR2_COTJA STANDARD; PRT: 1348 AA.
 AC P52583;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 2 PRECURSOR (EC 2.7.1.112)
 DE (VEGFR-2) (ENDOTHELIAL KINASE RECEPTOR EK1) (QUEK 1) (QUEK1).
 GN KDR OR FLK-1 OR EK1.

OS Coturnix coturnix japonica (Japanese quail).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Coturnix.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-EMBRYO;
 RX MEDLINE-97017121; PubMed-8863722;
 RA Eichmann A., Marcelle C., Breant C., Le Douarin N.M.;
 RT "Molecular cloning of Quek 1 and 2, two quail vascular endothelial
 RL growth factor (VEGF) receptor-like molecules.";
 RN Gene 174:3-8(1996).
 RP SEQUENCE OF 910-1348 FROM N.A.
 RC TISSUE-SPINAL CORD;
 RX MEDLINE-93378666; PubMed-8396413;
 RA Eichmann A., Marcelle C., Breant C., Le Douarin N.M.;
 RT "Two molecules related to the VEGF receptor are expressed in early
 RL endothelial cells during avian embryonic development.";
 RN Mech. Dev. 42:33-48(1993).
 RP SEQUENCE OF 764-880 FROM N.A., AND CHARACTERIZATION.
 RC TISSUE-EMBRYO;
 RX MEDLINE-95301109; PubMed-7781909;
 RA Flamme I., Breier G., Risau W.;
 RT "Vascular endothelial growth factor (VEGF) and VEGF receptor 2 (flk-1)
 RL are expressed during vasculogenesis and vascular differentiation in
 RL the quail embryo.";
 RN Dev. Biol. 169:1699-712(1995).
 CC -!- FUNCTION: RECEPTOR FOR VEGF OR VEGF-C. HAS A TYROSINE-PROTEIN
 CC KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM
 CC PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF
 CC VASCULAR PERMEABILITY.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: IN ALL ENDOTHELIAL TISSUES DURING ONSET OF
 CC VASCULARIZATION. IN LATER DEVELOPMENT, PRESENT IN LUNG, HEART,
 CC INTESTINE AND SKIN.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN WHOLE MESODERM AT ONSET OF
 CC GASTRULATION. FROM DAY 2, CONFINED TO ENDOTHELIAL TISSUES AND
 CC EXPRESSION CONTINUES TO BE WIDESPREAD THROUGHOUT VASCULARIZATION
 CC UNTIL E9 WHERE IT BECOMES RESTRICTED TO SPECIFIC REGIONS SUCH AS
 CC THE SPINAL CHORD AND HEART VALVES.
 CC -!- INDUCTION: IN VITRO, VEGF IS INDUCED BY BASIC FIBROBLAST GROWTH
 CC FACTOR (FGF), UNTOQUALLY IN THE FIRST 24 H OF CELL CULTURE.
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -!- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X83288; CAA38268.1; -;
 CC EMBL; S65205; AAB28127.1; -;
 CC EMBL; S78345; AAB34594.1; -;
 CC HSSP; P11362; IAGW;
 CC INTERPRO; IPR000719; -;
 CC INTERPRO; IPR001245; -;
 CC INTERPRO; IPR001824; -;
 CC INTERPRO; IPR003006; -;
 CC PFAM; PF00047; Ig; 6;
 CC PFAM; PF00069; pkinase; 2;
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1;
 CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1;
 CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1;
 CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1;
 CC Signal; Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; Phosphorylation; ATP-binding.
 FT SIGNAL 1 20
 FT POTENTIAL.

DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DE 01-OCT-2000 (Rel. 40, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 2 PRECURSOR (EC 2.7.1.112)
 DE (VEGFR-2) (KINASE INSERT DOMAIN RECEPTOR) (FLK-1).
 GS KDR OR FLK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yin L.Y., Wu Y., Patterson C.;
 RL "Full length human KDR/flk-1 sequence.";
 RN Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA YU Y., Whitney R.G., Sato J.D.;
 RL TISSUE-UMBILICAL VEIN;
 RN "Coding region for human VEGF receptor KDR (VEGFR-2).";
 RN Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 3-1356 FROM N.A.
 RC TISSUE-UMBILICAL VEIN;
 RX MEDLINE-92019839; PubMed-16563371;
 RA Terman B.I., Carrion M.E., Kovacs E., Rasmussen B.A., Eddy R.L.,
 RA Shows T.B.;
 RL "Identification of a new endothelial cell growth factor receptor
 tyrosine kinase.";
 RN Oncogene 6:1677-1683(1991).
 RN [4]
 RP SEQUENCE OF 1-22 FROM N.A.
 RX MEDLINE-96032749; PubMed-7559454;
 RA Patterson C., Perrella M.A., Hsieh C.M., Yoshizumi M., Lee M.E.,
 RA Harber E.;
 RL "Cloning and functional analysis of the promoter for KDR/flk-1, a
 receptor for vascular endothelial growth factor.";
 RN J. Biol. Chem. 270:23111-23118(1995).
 RN [5]
 RP FUNCTION.
 RX MEDLINE-93038639; PubMed-1417831;
 RA Terman B.I., Dougher-Vermaize M., Carrion M.E., Dimitrov D.,
 RA Armellino D.C., Gospodarowicz D., Boehlen P.;
 RL "Identification of the KDR tyrosine kinase as a receptor for vascular
 endothelial cell growth factor.";
 RN Biochem. Biophys. Res. Commun. 187:1579-1586(1992).
 CC -1- FUNCTION: RECEPTOR FOR VEGF OR VEGF-C. HAS A TYROSINE-PROTEIN
 KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM
 PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF
 VASCULAR PERMEABILITY.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC ENBL; AF035121; AAB88005.1; -
 CC ENBL; AF063658; AAC16450.1; -
 CC ENBL; X61655; CAA43837.1; -
 CC ENBL; L04947; AAA59459.1; -
 CC ENBL; X89776; CAA61916.1; -
 CC HSP; P11362; IAGW.
 CC MIN; I91306; -
 CC INTERPRO; IPR000719; -
 CC INTERPRO; IPR001245; -
 CC INTERPRO; IPR001824; -
 CC INTERPRO; IPR003006; -
 CC PFAM; PF00047; Ig; 6.

DR PFAM; PF00069; pkinase; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW Signal; Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.
 FT SIGNAL 1 19
 FT CHAIN 20 1356
 FT VASCULAR ENDOTHELIAL GROWTH FACTOR
 FT RECEPTOR 2.
 FT DOMAIN 20 764
 FT TRANSMEM 765 789
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IG-LIKE DOMAIN 1.
 FT DOMAIN 141 207
 FT IG-LIKE DOMAIN 2.
 FT DOMAIN 239 314
 FT IG-LIKE DOMAIN 3.
 FT DOMAIN 345 400
 FT IG-LIKE DOMAIN 4.
 FT DOMAIN 438 537
 FT IG-LIKE DOMAIN 5.
 FT DOMAIN 564 649
 FT IG-LIKE DOMAIN 6.
 FT DOMAIN 681 744
 FT IG-LIKE DOMAIN 7.
 FT DOMAIN 834 1162
 FT NP_BIND 840 848
 FT BINDING 868 868
 FT ACT_SITE 1028 1028
 FT BY SIMILARITY.
 FT CARBOHYD 46 46
 FT CARBOHYD 66 66
 FT CARBOHYD 96 96
 FT CARBOHYD 143 143
 FT CARBOHYD 158 158
 FT CARBOHYD 245 245
 FT CARBOHYD 318 318
 FT CARBOHYD 374 374
 FT CARBOHYD 395 395
 FT CARBOHYD 511 511
 FT CARBOHYD 523 523
 FT CARBOHYD 580 580
 FT CARBOHYD 613 613
 FT CARBOHYD 619 619
 FT CARBOHYD 631 631
 FT CARBOHYD 675 675
 FT CARBOHYD 704 704
 FT CARBOHYD 721 721
 FT MOD_RES 1059 1059
 FT CONFLICT 2 2
 FT CONFLICT 772 772
 FT CONFLICT 787 787
 FT CONFLICT 835 835
 FT CONFLICT 848 848
 FT CONFLICT 1347 1347
 FT SEQUENCE 1356 AA; 151526 MW; 59E7C4B05CFEBB3 CRC64;

Query Match 38.38; Score 2762; DB 1; Length 1356;
 Best Local Similarity 44.18; Pred. No. 6.1e-166;
 Matches 608; Conservative 203; Mismatches 480; Indels 88; Gaps 26;

QY 1 MORGAAALCLRLWLCGLLDGLVS--GYSMPTPTLNITEESHVIDTGSLSISCRGQHPLE 58
 DB 1 MQSKVLLAVLWLCVETRAASVGLPSVSLDPLRSIQKIDILTKANTLTQITCRGQRLD 60
 QY 59 NAWPGAQAPATGKDSQEDTGVVRC-EGTDARPYCKVLLLHVEHANDTGSVYCYIK 117
 DB 61 WLWPNNSG-----SEQRVEVTECSGDL----FCKLTIPKVIQNDTGAYKCFYR--- 106
 QY 118 ARIECTTAASSYVVRDFEQPFINKPDT-----LLVNRKDAWVPCVLSIPGLNVTI-- 169
 DB 107 ---ETDLASVIYVQDYSRSPFIASVSQHGVTYITENKNTVVPICLGSINLSICA 163
 QY 170 RSQSSVLPDQGEVWDDRRGLVSTPLLDALYLQETTGWQDQDFLSNPFLVHTGNEI 229
 DB 164 RYPEKRFVDPDGNRISWDSKSGFTIPSYMISYAGHVFCEAKINDESYQISIMYIVVVGTRI 223
 QY 230 YDIQLLPKSLLELVGKLVNCTVWAEFNSGVITFDWDYPCQKQAEKRGKWPERRSQOQTH 289


```

Db 1040 VRICDGLARDYKDPDVRKGRDARLPLKWAPEIFDRVYTIQSDVWSFGVLLWEIFSL 1099
QY 1111 GASPTGVOINEEFCORLDRGTMRAPELATPAIRIMLNCWSDGPKARPAFSELVEILG 1170
Db 1100 GASPTGPKVIDEEFCRLKEGTMRAPDYTPMYQTMLDCWHDENPORSFSELVEHLG 1159
QY 1171 DLQGRGLQEEVEECVAPRS-SQSEEGSFQSVSTMALHIAQADAEDSPPSLQRESLAAR 1229
Db 1160 NLLQNAOQDKDYILVPMSETLSMEEDSGLSPTSPSCWEEVEVCDP-----KFH 1211
QY 1230 YNNVWFFPGCLARGAETRGSRMKTFFEPFM-TPTTYKSGVDNQTDSGMVLASEEFPQIE 1288
Db 1212 YDNTAGISHYLQNSKRKSRPVSVKTFEDIPLEEPKVIPDDSTQDSGMVLASEELTLE 1271
QY 1289 SHROESGFCGKGPQNVAVTRAHPDSQGRRRRRPERGARGQVFNSE 1336
Db 1272 DRNKLSPSFG---GMPSKRSRESVASEGSNQ-----TSYGOSGYHSD 1310

RESULT 6
VGR2_RAT
ID VGR2_RAT STANDARD; PRT; 1343 AA.
AC 008775;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 2 PRECURSOR (EC 2.7.1.112)
DE (VEGFR-2) (PROTEIN-TYROSINE KINASE RECEPTOR FLK-1) (FETAL LIVER KINASE 1).
DE KDR OR FLK1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-RETINA:
RA Wen Y., Edelman J.L., De Vries G.W., Sachs G.:
RL Submitted (MAR-1997) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR VEGF OR VEGF-C. HAS A TYROSINE-PROTEIN
CC KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM
CC PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF
CC VASCULAR PERMEABILITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGFR RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U93306; AAB97508.1; -
DR EMBL; U93307; AAB97509.1; -
DR HSSP; P06239; 3LCK.
DR INTERPRO; IPR000719; -
DR INTERPRO; IPR001245; -
DR INTERPRO; IPR001824; -
DR INTERPRO; IPR003006; -
DR PFAM; PF00047; 1g; 6.
DR PFAM; PF00069; pkinase; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Signal; transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1343 VASCULAR ENDOTHELIAL GROWTH FACTOR
FT RECEPTOR 2.

```

```

FT DOMAIN 20 760 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 761 782 POTENTIAL.
FT DOMAIN 783 1343 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 46 109 IG-LIKE DOMAIN 1.
FT DOMAIN 141 207 IG-LIKE DOMAIN 2.
FT DOMAIN 239 304 IG-LIKE DOMAIN 3.
FT DOMAIN 345 400 IG-LIKE DOMAIN 4.
FT DOMAIN 438 533 IG-LIKE DOMAIN 5.
FT DOMAIN 560 645 IG-LIKE DOMAIN 6.
FT DOMAIN 677 740 IG-LIKE DOMAIN 7.
FT DOMAIN 830 1158 PROTEIN KINASE.
FT NP_BIND 836 844 ATP (BY SIMILARITY).
FT BINDING 864 864 ATP (BY SIMILARITY).
FT ACT_SITE 1024 1024 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 576 576 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 700 700 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MOD_RES 1055 1055 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 1343 AA; 150393 MW; AD7E509EB62D3FF4 CRC64;

Query Match 36.7%; Score 2646; DB 1; Length 1343;
Best Local Similarity 42.3%; Pred. No. 1.2e-158;
Matches 580; Conservative 210; Mismatches 484; Indels 96; Gaps 25;

QY 1 MORGAAALCLRLMCLGCLLDGLVS--GYSTPTPTLNTTESHVDTGDSLSISCRGHPLE 58
Db 1 MESRALLAVALFCVETRAASVGLPGDSLHPKLSQDKDILTLANTLIQIICRGORDLD 60
QY 59 WAPGQAQAPATGDKDSEDTGVVRDCEGTDAKYCKVLLHEVHANDTSYCYTYKAKA 118
Db 61 WLPNPT-----PRDSEERVLVTECGDS---IFCKTLTPRVVGVNDTGAYKCFYR--- 106
QY 119 RIEGTTAAS--YVVRDFEQPFINKPDI-----LLVNRKDMVPCVLSIPGLNVIL- 169
Db 107 ----DIDVSSIVYVQDRHSFFIASVDSDEHGIVITENKNTVTVIPCRGSISNLVSLC 162
QY 170 -RSQSVLWPDGOEVVMDRRGMLVSTPLLDHALYLCQETTWGDDQDFLSNPFVHTGNE 228
Db 163 ARYPKRRFPDGNRISWDSKGTIPSYMISYAGMVCEAKINDETYQSIMVILVVGVR 222
QY 229 LYDIQLLPKSLLELVGKLVLCNVAFNFFSVGVTFDWDYFGKQAKRGKWKVPERSSQOTH 288
Db 223 IDVVLSPPEHLELSAGEKLVLCNCTARTELNVGLDFSNQFPSSKHQHKIVNRDVKSLPG 282
QY 289 TELS---SILTIHNVSOHDLGSYVCKKANNQIQRFFRESTEVIVHENPFISVEMLKGPILA 345
Db 283 TVAKMFLSTLTIDSVTKSDQGEYTCAYSGMLTKNKIFVRVHTTPTFAFGSGMSILVEA 342
QY 346 TAGDELVLKPLVLAAYPPPEQWYEDKALSGRHS---PHALVLKEVTEASTGTTTALM 402
Db 343 TVGSG-VRIPVKILSYPADIKWENGRPIESNTYMIIVGDELTIMEVSEADAGNTVILT 401
QY 403 NSAAGLRNRNISLELVNVPVPPHKEAESSP-SIYSRHSRQALTCYAGVPLPLSIOWMR 461
Db 402 NPISMEKQSHMVSLLVNVVPPQIGERKALISPHDSYQYGTQMLTCTIVYANPLHHTWQ 461
QY 462 PWTPEKFAQSRLLRRQOQDLMPQCRDRAVTTQDAVNPISDLPWTVEFECKNKTYSKL 521
Db 462 LEEACSY-----RPSQTNPTCKEWRHVKDFQGNKIEVTNKOYALIEGKNKTSTVL 513

```



```
QY 522 VIONANVAMKYCVNKKYQODERLIYFYVTTIPDGTFTIESKPSEELLEQPVLLSCQAD 581
DB 514 VIOAAYVSAIKCEANKAGRGVRSFHVIRGPE---ITVOPATOPTERESMSLLCTAD 570
QY 582 SYKYEHLRWYRLNLSLTHDAHGPNLLLDCKNV---HLFATPLAASLEEVAPGARHATLS 637
DB 571 RNTFENLTYKLGSOATSVHMGESLTPVCKNLDALWKLNGTVFSNSTNDI-----LI 622
QY 638 LSTPRVAPHEGYCEVQDRSHDKHKKYSVQALAPRLTQHLTOLLVNVSDSLEM 697
DB 623 VAFONASLOQOGNYVCSAQDKTKRKHCLVKQLVILERMAMPITGNLENOTTIGETIEV 682
QY 698 QCLVAGAHAPSIVWYKDERLLEESQVDLADSNOKLSIQRVREEDAGRYLCSVCNAKGCV 757
DB 683 VCTSGNPTLITWFKDNETLVEDSGIVLKDGNRNLITIRVRKEDGGLTCAQCNVLGCA 742
QY 758 NSSASVAVGESDKSMEIVILVGTGVIAVFFWVLLLLIFCNMRRPAHADIKTYLSIIM 817
DB 743 RAETLFTIEGQEKTNLEVIILVGTAVIAMFFWLLVILVTRVKRANEGELKTGYLSIVM 802
QY 818 DPGEVPLEEOCEVLSYDASQWEPRELRHLGRVLGYGAGSKVVEASAFGIHKSSCDTYA 877
DB 803 DPDELPLDERCELRPLDASQWEPRELRHLGRVLGYGAGSKVVEASAFGIHKSSCDTYA 862
QY 878 VKMLKEGATASEHRLMSLKIILHIGNHLNVNLLGACTKPGQPLMWIVFECKYGNLSN 937
DB 863 VKMLKEGATASEHRLMSLKIILHIGNHLNVNLLGACTKPGQPLMWIVFECKYGNLSN 922
QY 938 FLAKRADFSQAEKSPQGRFRA---MWELARLDRRPGSSDRVLFAFSKTEGGAR 993
DB 923 YLRGRKNEFVYKSG---AFRSKDYVGLS-VDLKR--RLDSITSSQSSASGFFE 975
QY 994 RAS-----PDQEAEDLWLSPLTMDLVYCSFQVARGMEFLASRKCITRDLAARNILSES 1048
DB 976 EKSLDVEEAESELYKDFLEHLICYSFQVAKMEFLASRKCITRDLAARNILSEK 1035
QY 1049 DVYKICDFGLARDIYKDPDYVRKGSARLPKLMWAPESIFDKVYTTQSDVWSFGVLLWEIF 1108
DB 1036 NVYKICDFGLARDIYKDPDYVRKGSARLPKLMWAPETIFDRIYTIQSGVWSFGVLLWEIF 1095
QY 1109 SLGASVPGVQINEEFCORLRDGTMRAPELATPATRRIMLNCWSGDPKARPAFSELVEI 1168
DB 1096 SLGASVPGVQINEEFCORLRDGTMRAPDYTPPYMTQMLDCWHDNPORPAFSELVEH 1155
QY 1169 LGLDQGRGLQEEVEECMAPRS--SQSSESGFSQVSTMALHTAQADAEDSPPSLQARHSLA 1227
DB 1156 LGLDQGRGLQEEVEECMAPRS--SQSSESGFSQVSTMALHTAQADAEDSPPSLQARHSLA 1207
QY 1228 ARYNNVSVFPGCLARGAETRGSRMKTFFEPFM-TPTTYKGSVDNQTDSQVNLASEFEQ 1286
DB 1208 FHYDNTAGISHYLQNSKRSPVSVKTFEDIPLEEPKVPYIPDQSDTDSQVNLASELKT 1267
QY 1287 IESRHQESGFSCKGPGQNVAVTRAHPDQSGRRRRRPERGARGGOVYNSE 1336
DB 1268 LEDNKLSPSPG---GMPPSKRSRESVASEGSNQ-----TSGIQSGYHSD 1308
RESULT 7
VGR1_MOUSE
ID VGR1_MOUSE STANDARD; PRT: 1333 AA.
AC P35969; Q61517; Q55094;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
DE (VEGFR-1) (TYROSINE-PROTEIN KINASE RECEPTOR FLT) (FLT-1) (EMBRYONIC
DE RECEPTOR KINASE 2).
GN FLT1 OR FLT OR EMBK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
```


FT NP_BIND 834 842 ATP (BY SIMILARITY).
 FT BINDING 862 862 ATP (BY SIMILARITY).
 FT ACT_SITE 1022 1022 BY SIMILARITY.
 FT MOD_RES 1053 1053 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT DISULFID 53 108 POTENTIAL.
 FT DISULFID 159 208 POTENTIAL.
 FT DISULFID 253 312 POTENTIAL.
 FT DISULFID 455 536 POTENTIAL.
 FT DISULFID 578 637 POTENTIAL.
 FT DISULFID 683 732 POTENTIAL.
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 517 517 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 598 598 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 626 626 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 667 667 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 714 714 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 158 158 MISSING (IN REF. 2).
 FT CONFLICT 211 211 MISSING (IN REF. 2).
 FT CONFLICT 245 245 H -> L (IN REF. 2).
 FT CONFLICT 603 603 H -> N (IN REF. 2).
 FT CONFLICT 609 615 KMATOD -> NGHSS (IN REF. 2).
 FT CONFLICT 696 696 F -> L (IN REF. 2).
 FT CONFLICT 734 734 A -> S (IN REF. 2).
 FT CONFLICT 765 765 C -> Y (IN REF. 2).
 FT CONFLICT 820 820 K -> N (IN REF. 2).
 FT CONFLICT 1009 1009 G -> R (IN REF. 2).
 FT CONFLICT 1181 1181 S -> G (IN REF. 3).
 FT CONFLICT 1181 1181 S -> N (IN REF. 2).
 FT CONFLICT 1193 1194 LF -> RG (IN REF. 2).
 FT CONFLICT 1279 1279 KS -> PR (IN REF. 2).
 SQ SEQUENCE 1333 AA; 149875 MW; C06533B7ECB404C CRC64;

Query Match 34.28; Score 2469; DB 1; Length 1333;
 Best Local Similarity 42.48; Pred. No. 1.6e-147; Indels 96; Gaps 27;
 Matches 555; Conservative 186; Mismatches 472;

QY 5 AALCLRLWCLGLDGLVSGYSMTPTTINITEESHVIDTGDLSLSCRGQHPLEWAPGA 64
 DB 8 AVLPPALLGCL-LLTGYSGSKLVPELSLAGTQHVQAGQTLFLKCRGEAAHWSLP-- 64
 QY 65 QAPATGDKDSEDGTVWRDCGTDAKYCKVLLHEVHNDTGSYVCYKIKARIEGTT 124
 DB 65 ---TTVSQEDKRLSTPPSACGRNQCSTLTLDTAQNHGTGLTYCRYLPTSTSKKKA 121
 QY 125 AASSVVFVRDFEQPINK---PDTLLVNRKDAMVPCVLSIPGLNVTLSR-QSSVLMPD 179
 DB 122 ESSVIFVSDAGSPFIEHMTDIPKLHVHTEGRQLIICRVTSNPVTVLKKEPDTLTPD 181
 QY 180 GQEVWDDRRGMLVSTPLLDALYLQCTTWGQDDFLSNPFLVHTGNELYDIQLLPKRS 239
 DB 182 GQRTWDSRRGFIAANATYKKEIGLNCATVNGHLYQTN-YLTHRQNTILDVQIRPSP 240
 QY 240 LELLVGEKLVNCTWAEFSGVTFDWDYCKQAEGRGWVPER--RSOOTHTESSILTI 297
 DB 241 VRLHGQTLVLNCTATTENLRVQSNWNPYKATKRAS-IRQIDRSHSHNNVHFSLVKI 299
 QY 298 HNVSDRLGSLVCKANNGIQRSTEVIVHENPFISVWELKGPTEAGDELVKLPVK 357
 DB 300 NVNVEDRGLVTCRVKSGSFSQSFNTSVHYVEKGFISVKHKKQPVQVETAGRRSVLSMK 359
 QY 358 LAAYPPPEFQWKDGKAL---SGRHSR--ALVLKEVTEASTGYTTLALNLSAAGLRNI 412
 DB 360 VRAFPSPVLWLDKSPATLSARLYVHGYSLLIKDKVTTEDAGDYITLLGFKQSLFKNL 419
 QY 413 SLELVNVPVPOIHEKAS---SPSTYSRHSRQALTCTAYGVLPLPSIQHWHRPMTCKMF 469

DB 420 TATLIVNKPQIYKSVSSLPPLYPGLSGRQVLTCTVYIGPRP-TITLWH----PC---- 472
 QY 470 AQBSLRRROODLMPQCRD-WRAVITQDAVNPIESLDTWTTEVEGKNKYSKLVIQANVY 528
 DB 473 --HHNHSKERYDCTENEESFTLDPSSNIGNRIESQRTVIEGNTKTVSLVWADSTQ 530
 QY 529 SAMYKCVSNKYGQDERLIYVVTIPDGF--TIESKPSEELLEGQPVLLSQADSKYKE 586
 DB 531 PGYISCRARFNKGTVERNIKFYVDVPGNFHVSLEKMPA---EGEDLKSCVWNKFLR 586
 QY 587 HLRWYRL---NLSLHDAHGNPLLLDCKNWHFATPLAASLEEVAPGAHATLSISIPRV 643
 DB 587 DITWILLRTVNNRTH-----HSISKQKMATQDYS-----ITLNLVIRNV 627
 QY 644 APEHEGHHYCEVDQRSHDKHCKKYSVQALEAPLTONLTDLLVNVSDSLEMCCLVAG 703
 DB 628 SLEDSGTACRAHNIYTGEDILRKTEVLVRDSEAPHLQNLSDYEVISGSTLDCQARG 687
 QY 704 AHAPSTVWYKDERLLEEKSGVDLADSNOKLSQVRVEDAGRYLCVCAKAGCVNNSASV 763
 DB 688 VPAQITWFKNNHKIQQEPGIIILGPGNSTLFIERTVEDEGVYRCRATNKGAVESAAYL 747
 QY 764 AVEGSDKSGMEIVILVGVGVIAVFWVLLLLIFCNRHRAHADIKTGILSIIMDPEVP 833
 DB 748 TVQGTSDKSNLELITLTCVAATLEWLLTLFIRKLKR-SSSEVKTDLISIIMDPEVP 806
 QY 824 LEEQCEVLSYDASOWEFPFRERLHGRVIGVAGFGKVVVEASAFGIHKGSSCDTVAVKMLKE 883
 DB 807 LDEQCERLPYDASKWFAERLKLKSLGSLGAGFGKVVQASAFGIKASPTCRTVAVKMLKE 866
 QY 884 GATASEHRLMSELKILIHGNHNLVNNLLGACTRPOGPLWIVIVEFCYKGNLSNFLRAKR 943
 DB 867 GATASEVKALMTLTKILTHIGHHLNVNLLGACTKOGGPLWIVIVECYKGNLSNLYLKKR 926
 QY 944 DAFSPCAEKSP-----QGRFRPAVVELARLDRRRGSSDRVLFARF-----SKTEGGA 992
 DB 927 DLF--CLNKDAALHMLKESLEPGLEQOGKPRLDSSVSSSTSSSPEDRSVSDVEG-- 982
 QY 993 RRASPOQEAEDLWLSPLTMEDLVCSYFQVARGMEFLASRKCIRHDLAARNILISESDVVK 1052
 DB 983 -----DEDYSEISKQPLTMEDLISYFQVARGMEFLSSRKCIRHDLAARNILISENNVK 1037
 QY 1053 ICDFGLARDIYKDPDVKYRGKSARLPLKWNAPESIFDKVYTTOSDVMSFGVLLWEIFSLGA 1112
 DB 1038 ICDFGLARDIYKDPVYVRGDTRLPLKWNAPESIFDKVYSTKSDVMSYGVLLWEIFSLGG 1097
 QY 1113 SPYPGVQINEFCORLDCRTRAPELATPAIRRMILCMWCGDKPAPAFSELVEILGDL 1172
 DB 1098 SPIPGVOMDEDFCSRLKEGHRMTPEYATPEIYQIMLDCWHKDKPKEPRFAELVEKLGDL 1157
 QY 1173 LQGRGLQEEVEVCHMAPRSSQSESGFSQVSTMALHIAQADSPSPSLQRHSLAARYN 1232
 DB 1158 LQANVQDQGD--YIPLNALITRNSSTYPTFSEDLFKDGFADPHFHSGSSDDVRYVN 1215
 QY 1233 WVSFPGCLARGAETRRSSRMKTEEPFMPPTTYKGSVDNQTDSCMVLAS 1281
 DB 1216 AFKF-----MSLERIKTFEE--LSPNSTSMFEDYQDSTLLGS 1252

RESULT 8
 VGR1_HUMAN
 ID VGR1_HUMAN STANDARD; PRT; 1338 AA.
 AC P17948; P16057; O60722;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE VASCULAR ENDOTHELIAL GROWTH FACTOR RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
 DE (VEGFR-1) (TYROSINE-PROTEIN KINASE RECEPTOR FLT) (FLT-1) (TYROSINE-
 DE PROTEIN KINASE FRT).
 GN FLT1 OR FLT OR FRT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

[1]
SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RX MEDLINE=90221591; PubMed-2158038;
RA Shibuya M., Yamaguchi S., Yamane A., Ikeda T., Tojo A.,
RA Matsushime H., Sato M.;
RT "Nucleotide sequence and expression of a novel human receptor-type
RT tyrosine kinase gene (flt) closely related to the fms family.";
RL Oncogene 5:519-524(1990).
[2]
SEQUENCE FROM N.A.
RP TISSUE=UMBILICAL VEIN;
RA Yu Y., Whitney R.G., Sato J.D.;
RT "Coding region for human VEGF receptor FLT1 (VEGFR-1).";
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE OF 1018-1058 FROM N.A.
RX MEDLINE=87307638; PubMed-3040650;
RA Matsushime H., Yoshida M.C., Sasaki M., Shibuya M.;
RT "A possible new member of tyrosine kinase family, human frt sequence,
RT is highly conserved in vertebrates and located on human chromosome
RT 13.";
RL Jpn. J. Cancer Res. 78:655-661(1987).
CC -1- FUNCTION: RECEPTOR FOR VEGF AND PLGF. HAS A TYROSINE-PROTEIN
CC KINASE ACTIVITY. THE VEGF-KINASE LIGAND/RECEPTOR SIGNALING SYSTEM
CC PLAYS A KEY ROLE IN VASCULAR DEVELOPMENT AND REGULATION OF
CC VASCULAR PERMEABILITY.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: MOSTLY IN NORMAL LUNG, BUT ALSO IN PLACENTA,
CC LIVER, KIDNEY, HEART AND BRAIN TISSUES. IT IS NOT EXPRESSED IN
CC TUMOR CELL LINES.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE DOMAINS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X51602; CAA35946.1; -
EMBL; AF063657; AAC16449.1; -
EMBL; D00133; BAA00080.1; -
PIR; S09982; S09982.
HSSP; P11362; 1FG1.
MIN; 165070; -
INTERPRO; IPR000719; -
INTERPRO; IPR001245; -
INTERPRO; IPR001824; -
INTERPRO; IPR003006; -
PFAM; PF00047; 1g; 7.
PFAM; PF00069; pkinase; 2.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
Receptor; Tyrosine-protein kinase; Phosphorylation; ATP-binding;
Signal; Transmembrane; Signal; Immunoglobulin domain; Glycoprotein.
SIGNAL 1 22
CHAIN 23 1338
VASCULAR ENDOTHELIAL GROWTH FACTOR
RECEPTOR 1.
EXTRACELLULAR (POTENTIAL).
DOMAIN 23 758
TRANSMEM 759 780
POTENTIAL.
DOMAIN 781 1338
CYTOPLASMIC (POTENTIAL).
DOMAIN 46 114
IG-LIKE DOMAIN 1.
DOMAIN 151 214
IG-LIKE DOMAIN 2.
DOMAIN 245 318
IG-LIKE DOMAIN 3.
DOMAIN 363 404
IG-LIKE DOMAIN 4.
DOMAIN 447 542
IG-LIKE DOMAIN 5.
DOMAIN 570 643
IG-LIKE DOMAIN 6.

FT	DOMAIN	675	738	IG-LIKE DOMAIN 7.
FT	DOMAIN	827	1158	PROTEIN KINASE.
FT	NP_BIND	833	841	ATP (BY SIMILARITY).
FT	BINDING	861	861	ATP (BY SIMILARITY).
FT	ACT_SITE	1022	1022	BY SIMILARITY.
FT	MOD_RES	1053	1053	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT	DISULFID	53	107	POTENTIAL.
FT	DISULFID	158	207	POTENTIAL.
FT	DISULFID	252	311	POTENTIAL.
FT	DISULFID	454	535	POTENTIAL.
FT	DISULFID	577	636	POTENTIAL.
FT	DISULFID	682	731	POTENTIAL.
FT	CARBOHYD	100	100	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	164	164	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	196	196	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	251	251	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	323	323	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	402	402	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	417	417	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	474	474	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	547	547	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	597	597	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	620	620	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	625	625	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	666	666	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT	779	779	L -> F (IN REF. 2).
SQ	SEQUENCE	1338	AA; 150733	MW; 54F1554E5590787F
CRC64;				
Query Match 34.2%; Score 2463; DB 1; Length 1338;				
Best Local Similarity 39.9%; Pred. No. 3.8e-147;				
Matches 547; Conservative 228; Mismatches 469; Indels 138; Gaps 28;				
QY	4	GAALCLRLWLCGLDGLVSGYSMTPTPLNITEESHVDTGDSLSISCRGQHPLEWAPG	63	
Db	8	GVLLC-ALLSCL-LLTGSSSGSKLMDPELSLKGTOHMQAGOTLQCRGEAAHKS---	62	
QY	64	AQEAATGDKDSEDGVVYRDCGTDAAPVCKVLLHVEHANDTGSVCYVYKIKARIGT	123	
Db	63	---LPEVSKESERLSIKSACGRNGKFCSTLTLTAAQNHGTGYSCKYLAVPSTKKE	119	
QY	124	TAASSYVFVDFEQPFINK-----PDTLVNRKDMWPCVLSIPGLNLTLS-QSVMWP	178	
Db	120	TESATYIFISDTGRPEVEMYSEIPEIHTTEGRELVIPOCVTSPTNITVTKKFLDTLP	179	
QY	179	DQGEVVDORGLVSTPLLDALYLQCTTGWDDFLSNPFLVHTGNYDIOLLPK	238	
Db	180	DGKRIIWSRKGFIISNATYKEIGLTCATVNGHLYKTN-YLTHROTWTIDVQLSTPR	238	
QY	239	SLELVGKLVNCTVMAEFNSGVTFDNDYPOQAERGKWPERRSQOHTELS---SIL	295	
Db	239	PVKLLRGHTLVNCTATTPLNTRVQMTWSYPDEKNKRAS--VRRRIDQSNHANFYSL	296	
QY	296	THNVSHDLGSLVCKANNGIQRFRETEVIVHENPFIENVHKGPILEATAGDELVKLP	355	
Db	297	TIDMQNKDGLYTCRVRSGPSFKSVNTSVHYDKAFITVKHKKQOVLETVAGKRSRLS	356	
QY	356	VKLAAYPPEFQYKDKGKALSGRHS-----PHALVKEVTEASTGTITLALNNSAAGLR	410	
Db	357	MVKAFPSPEVWLDGLPATEKSARYLARGYSLLIKDVEEDAGNTILLSIKQSNVFK	416	
QY	411	NISLELVNVPPOIHEKEASS---PSIYSRHSQALCTAYGVPLPLSTOWHWRPWPCK	467	
Db	417	LTATLVNKPQIYKEAVSSFPDPAIPLGSRQILTCYAYGIPQ-P-TIKWFPHPCHNH	475	
QY	468	MFAQSLRRRQQDLMPCQCRDMRAVTTQDAVNPIESLDTWTEFEVGNKNTVSKLVIONAN	527	
Db	476	SEARCDSCSNNEESFILD-----ADSNMGRNIESITORMAIECKNMASTLVVADSR	528	
QY	528	VSAMYKCVSNKVGQDERLIYFVTTIPDGF--TIESKSEELLEGPVLLSQADSYKY	565	
Db	529	ISGIYICIASNKGTVGRNISFYITDPVNGFHVNLKQHTP-----EGEDLKSCTVKNFLY	584	

01-NOV-1997 (Rel. 35, Last annotation update)
DE ALPHA PLATELET-DERIVED GROWTH FACTOR PRECURSOR (EC 2.7.1.112)
GN (PDGF-R-ALPHA).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP MEDLINE-91061789; PubMed-2174116;
RA Stiles C.D., Wang C.;
RT "Retinoic acid promotes transcription of the platelet-derived growth
RL factor alpha-receptor gene.";
Mol. Cell. Biol. 10:6781-6784(1990).
[2]
RP MEDLINE-92334866; PubMed-1321404;
RA Do M.S., Fitzer-Attas C., Gubbay J., Greenfield L., Feldman M.,
RA Eisenbach L.;
RT "Mouse platelet-derived growth factor alpha receptor: sequence,
RT tissue-specific expression and correlation with metastatic
RT phenotype".
Oncogene 7:1567-1575(1992).
RL FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND
CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR CAN BIND
CC EITHER PDGF-A OR PDGF-B.
CC -1- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA
CC SUBUNITS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
EMBL: M84607; AAA39904.1; -;
DR EMBL: M57683; AAA39733.1; -;
DR PIR: S33727; S33727.
DR HSP; P11362; IAGW.
DR MGD; MGI:97530; PDGFRA.
DR INTERPRO: IPR000719; -;
DR INTERPRO: IPR001245; -;
DR INTERPRO: IPR001824; -;
DR INTERPRO: IPR003006; -;
DR PFAM: PF00047; Ig; 4.
DR PFAM: PF00049; pkinase; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Tyrosine-protein kinase; Receptor; Transmembrane; Glycoprotein;
KW transferase; Phosphorylation; ATP-binding; Immunoglobulin domain;
KW Signal.
FT SIGNAL. 1 24
FT CHAIN 25 1089
FT DOMAIN 25 525 ALPHA PLATELET-DERIVED GROWTH FACTOR
FT TRANSMEM 526 549 RECEPTOR.
FT DOMAIN 550 1089 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 593 954 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 599 607 PROTEIN KINASE.
FT BINDING 627 627 ATP (BY SIMILARITY).
FT ACT_SITE 818 818 ATP (BY SIMILARITY).
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 359 359 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 506 506 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 192 192 E -> A (IN REF. 1).
FT CONFLICT 202 202 E -> A (IN REF. 1).
FT CONFLICT 252 252 L -> V (IN REF. 1).
FT CONFLICT 271 271 G -> S (IN REF. 1).
FT CONFLICT 322 322 A -> P (IN REF. 1).
FT CONFLICT 326 326 GT -> EG (IN REF. 1).
FT CONFLICT 439 440 A -> E (IN REF. 1).
FT CONFLICT 529 529 A -> D (IN REF. 1).
FT CONFLICT 737 737 Y -> D (IN REF. 1).
FT CONFLICT 849 849 E -> D (IN REF. 1).
FT CONFLICT 936 936 V -> L (IN REF. 1).
FT CONFLICT 950 950 S -> G (IN REF. 1).
FT CONFLICT 1005 1005 S -> G (IN REF. 1).
SQ SEQUENCE 1089 AA; 122727 MW; B3A3B246716249E0 CRC64;
Query Match 16.1%; Score 1160; DB 1; Length 1089;
Best Local Similarity 28.4%; Pred. No. 2.6e-65;
Matches 354; Conservative 164; Mismatches 358; Indels 372; Gaps 44;
QY 14 CLGLDLGLVGYSTPPTLNTTESHVDTGDSLSISCRQHPLHNAWPGQAQAPATGDK 73
DB 13 CLLTGPGILSCQLLPILP-NENEKIVQLNSSFSLRCVGESEVSMOHPMSE----- 64
QY 74 DSEDTGVVRCDEGTDPYCKVLLHEVHANTDGSVYVYKYI---KARTEGTTAASSV 130
DB 65 --EDPNVEIRSEENNSGLFTVLEVNAASAHGWTCTYNNHTQDSEIEG---RHII 119
QY 131 FVRDFEQPFI--NRPDPL-LVNRKDAWVPCVLSIPGLNVTLSQSSVLPDQGVVDD 187
DB 120 YVPDPMFAVPLGMDSLVIVEEDDSAIIPCTDPTQVTLHNGRLV-----PASYS 174
QY 188 RRGVLSTPLLDALYLQCTTNGDQDFLNPFLVHITGNELYDIOLPKRSLLE----- 242
DB 175 RQGF---NGTSVGPYI-CENTVGRTEKISEP-----NVTALKATSELNEMDARQT 223
QY 243 --LVGEKLVNCTVMAEFNSGTFDWDYPGQKQERKWPERRSQOHTIELSSI-----L 295
DB 224 VKAGETIVTCAVEN-NEVVDLQWTYPGEVRNKGITMLE-----EIKLPSIKLVTL 275
QY 296 TIHNVSQHDLGSLVCKKANGLORETEVI--VHENPFISVEHKLGPILLEATAGDELVK 353
DB 276 TVPKATVKDSGEYECARQATKEVKEMKRVTSVHEKGFVEIETPTGO-LEAVNLHVE 334
QY 354 LPVKLAAYPPPEFOWYKDG-----KALSGRHSPLHLVLEKVEATEASTGTYY 398
DB 335 FVVEQAYPTPRISWLKDNLTLENTITTDVQKSOETRYQ-SKLKLIRAKEEDSGHYT 393
QY 399 LALNWSAAGLRNLSLELVNVPQIHEKEASSIYSRH-----SRQALTCATGVPLPL 454
DB 394 IIVQNEEDD--VKSYTFELSTLVP-----ASILDVDDHSGGGGTQVRCAGTGLPE 444
QY 455 SIQWHRRWPTPCMKFAQRSLRRROODLMPQCRDWRVTTQDAVNPIESLDT-WTEFEV 513
DB 445 I-----DMICKH-----IKKCN-----DTSMT----- 463
QY 514 KNTVSKLVQIONANSYKCVSNKVGODERLIYFYVTTP--DGTIESKPELELEG 571
DB 464 -----VLASVNSI-----ITELPRGRSTVEGRVSFAKYE- 494
QY 572 QPVLLSCQADSYKYEHLRWYRLNLTSLDHAHGNPLLDCKKNVHLFATPLAASLEEVAPGA 631
DB 495 ETIAVRCLAKN-----NLSV-----VARELKLVAPTL 521
QY 632 RHATLSLSTPRVAPEHGHYCEVOODRSHDKHCKHKYLSVOALEAPRLTOMLTOLLVNV 691
DB 522 R-----SELTAAAVLVLLVTVIVSLVLV 546


```

Db 224 YKAGETIVTCAVFN--NEVDLQWYGEVNRKNGITMLE-----EIKLSIKLVYILT 275
QY 297 IHNVSQDLGSLGYCKANRGIQFRE--STEVIVHENPFISVEMWLGKPILEATAGDELKVL 354
Db 276 VPKATVYSGDYECARQATKEVKEMKTVTISVHEKGFVOIRPTFGH-LETVNLHQVREF 334
QY 355 PVKLAAYPPPEFYWDYKDGKALSGRSHFALVLKEVT-----EAS 393
Db 335 VWEQAYPTPRISWLNKDNLTLEN-----LTEITDVORSQETRYOSKLIKIRAKEED 387
QY 394 TGYTTLALWNSAAGLRNLSLELVNVPQIHEKEASSPSIYSRHSR---QALTCATYGV 450
Db 388 SGHYTIIVQNDND--MKSITFELSTLVPASILE-----LVDDHGGGGQIVRCTAEGT 439
QY 451 PLPLSTOMWRPWPCKMFAQRSLRRQQODLMPQCRD---WRVTTQDVNPNIESLTDW 507
Db 440 PLP-NIE-----WMICKD-----IKKCNNDTSM-TVLASVSNII-----472
QY 508 TEFVE-GKNKTVSKLVIONANVSAMKCVVSNKVGQDERLIIFYVTTPDGFTIESKPSE 566
Db 473 TEFHQRGRSVTEGRVSFAKVEETIAVRCLAKNDLIGNR-----511
QY 567 ELLEGQVLLSCQADSKYKYEHLRWYRLNLSLHDAGNPLLLDCKNHLFATPLAASLEE 626
Db 512 -----VIWK-----OKPRYEIRWRV 560
QY 747 LCVNNAKGCNVSSASVAVESGDSKGSMEITVLVGTGVIAVFEFWLLLLIFCNMRPAHA 806
Db 561 IESI-----SPD-----567
QY 807 DIKTGYLSIIMDPEVPLEQCEYLSYDASQWEPFRERHLGRVLGYGAFKVKVEASAFG 866
Db 568 ---GHEYIVDPMPQLP-----YD-SRWEFFRDLGLGRILGSGAFKGVVEGTAYG 613
QY 867 IHKSSCDTVAVKMLKESATASEHRLMSKLTILHIGNHLNVNLLGACTKPGQPLMWI 926
Db 614 LRSQPVKVAVMKLTARSEKQALMSKLTIMHLGPHLNVNLLGACTK-SGFIYII 672
QY 927 VEFCKYGNLSNFKRAKDAF-----SPCAEK-----SPEQRGRFRA 962
Db 673 TEYCFYGLVNLHKNRDSFMSRHPKPKDLDIFGLNPADESTRSVYLSFENNNGDYVD 732
QY 963 MVE-----LARDRRRPGSDRVLFARFSEKTEGGARRASPDQAEEDLWSP-----LTM 1011
Db 733 MKQADTQYVPMLEKVEKYSYDQISLYDRPASYKKKMSLDSEAKML-LSDDOSEGLTL 791
QY 1012 EDLVYCSFQVARGMEFLASRCKIHRDLAARNILLSESDVVKICDFGLARDIYKDPDYVRK 1071
Db 792 LDLLSETYQVARGMEFLASKNCVHRDLAARNVLLAQCKIVKICDFGLARDIMHDSNYSVK 851
QY 1072 GSARPLKWAPEISFDKYVTTQSDVNSFGVLLWEIFSLGASPYPGVQINEECORLQD 1131
Db 852 GSTPLPVKWAPEISFDNLTYSYVGLLWEIFSLGDTYPYPMHVDSTYFNKISG 911
QY 1132 TRMRAPELATPAIRIMLNCWSGDKPARKAPAFSELVEITGLDLOGRGLQEEECYCMAPRSS 1191
Db 912 YRMKAPDQATSEVYIEMVQCNSEPEKRPSPYHLSEIVENLLPGQYKKSVEKIHLL-----966
QY 1192 OSSEGSFSQVSTWALHTAQADAEDSPSLQRHSLARYNWNYSFPCLARGAETGSSR 1251
Db 967 -----DFLKSDDHPAVARNRVDSNA-----YIGVTYRN-----EEDK 998
QY 1252 MKTFFEFPTPTTYKGSVDNQ---TDSQWVL-----ASEEFEQIESRHRQES 1295

```

```

Db 999 LKWE-----GGLDEORLSADSGYIIPLDIPDPVEEDLGKRRHSS 1041

RESULT 13
PGDS_XENLA STANDARD; PRT: 1087 AA.
ID PGDS_XENLA
AC P26619;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALPHA PLATELET-DERIVED GROWTH FACTOR RECEPTOR PRECURSOR (BC 2.7.1.112)
OS (PDGF-R-ALPHA).
DE Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93365089; PubMed-8358864;
RA Jones S.D., Ho L., Smith J.C., Yordan C., Stiles C.D., Mercola M.;
RT "The Xenopus platelet-derived growth factor alpha receptor: cDNA
RT cloning and demonstration that mesoderm induction establishes the
RT lineage-specific pattern of ligand and receptor gene expression.";
RL Dev. Genet. 14:185-193(1993)
CC CC -!- FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND
CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR CAN BIND
CC EITHER PDGF-A OR PDGF-B.
CC -!- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA
CC SUBUNITS.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M80798; AAA49929.1; --
DR HSP: P11362; IAGW.
DR INTERPRO: IPR000719; --
DR INTERPRO: IPR001245; --
DR INTERPRO: IPR001824; --
DR INTERPRO: IPR003006; --
DR PFAM: PF00047; Ig; 3.
DR PFAM: PF00069; pkinase; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
KW Tyrosine-protein kinase; Receptor; transmembrane; Glycoprotein;
KW transferase; Phosphorylation; ATP-binding; Immunoglobulin domain;
KW Signal.
FT SIGNAL 1 ?
FT CHAIN 1 ?
FT ? ALPHA PLATELET-DERIVED GROWTH FACTOR
FT RECEPTOR
FT EXTRACELLULAR (POTENTIAL).
FT DOMAIN 529 ? 528
FT TRANSMEM 552 1087
FT DOMAIN 595 970
FT DOMAIN 601 609
FT NP_BIND 629 629
FT BINDING 629 629
FT ACT_SITE 818 818
FT MOD_RES 849 849
FT CARBOHYD 77 77
FT CARBOHYD 104 104
FT CARBOHYD 216 216
FT CARBOHYD 282 282
FT CARBOHYD 309 309
FT CARBOHYD 309

```


QY 482 LMPQCRDRAVTTQDAVNPI-----ESLDT-----MTEFVEGKNKTVSKLVIONANVSA 530
 Db 448 --STCDLRCRCPKSPPLGNSKEESQLENTNFTFEE--DOEYEVYSTLRHRVQPL 503
 QY 531 MYKCVVSNKVGODERLIYFYVTTPDGTIESKPEELLEGOPVLLSCQADSYKYEHLRW 590
 Db 504 SVRCMLQNSMGDSQ-----EVTVP----- 524
 QY 591 YRLNSTLHDANGNPLLLDCRNVHLFATPLAASLEEVAPGARHATLSLSIPRVAPEHEGH 650
 Db 525 ----- 524
 QY 651 YVEVQDRSHDKHCKKYLVSQALEAPRLATQNLTDLVNVSDSLEMQLVAGAHAPSIV 710
 Db 525 -----HSLPFKVVVISA----- 536
 QY 711 WYKDERLLEKSGVDLADSNQKLSIQRYVEDAGRYLCVCNAKGCYNSSASVAVEGSED 770
 Db 537 ----- 536
 QY 771 KGSMEIVILVGTGVIAGVFWVLLLLIFCNMRRPAHADIK-----TGYLSIIMDPGE 821
 Db 537 -----ILALVLTVIS-----LILILMLWQKKPY-EIRWKVIESVSDGHEIYIVDPVQ 585
 QY 822 VPLEOCEYLSYDASQWEPREPRHLGRVLGYGAFGKVVEASAFGIHKGSSCDTVAVKML 881
 Db 586 LP-----YD-STWELPRDQVLGRTLGSAGFQVVEATAHGLSHSQATMKVAVKML 635
 QY 882 KEGATASEHRALMSELKILIHIGNHLNVNLLGACTKQGPPLMWIVFECKYGNLSNLR 941
 Db 636 KSTARSEKQALMSELKINSHLGNVNLGACTK-GGPYIIITECRYGDLVDYLHR 694
 QY 942 KRDAFSPCAEKSPQGRFRAMVELARLDRRRPGSSD---RYLFARFS-----KTE 989
 Db 695 NKHTFL-----QR-----HSNKHCPSPAELYSNALPVGFSLSPLNLTGESD 736
 QY 990 GGARRASPOEAE-----DL-----WLSP----- 1008
 Db 737 GGYMNSKDESIDYVPMIDMKDGIADIESFSPYADNYVPSAPERTYRATLINDSPV 796
 QY 1009 LTMEDIVCYSFQVARGMEFLASRCKIHRDLAARNILLSESDVVKICDFGLARDIYKDPDY 1068
 Db 797 LSYTDLVGFYSQVANGMDFLASKNCVHRDLAARNVLICEGKLVKICDFGLARDINRDSNY 856
 QY 1069 VRKGSARLPLKWNAPESIFDKVYTTQSDVWFGVLLWEIFSLGASPPYGVQINEEFQORL 1128
 Db 857 ISKGSTYLPKWNAPESIFNSLYTTLSDVWFGVLLWEIFSLGASPPYGVQINEEFQORL 916
 QY 1129 RQCTRMRAPELATPATRRIMLNCWSDPKARPAFSELVEILGDLQGRGL-----QEEEE 1183
 Db 917 KRGYRMAQPAHASDEIYEMQKCEKEETRPFPFQVLLVLLERLL-GEYKXKKYQOVDDEE 975
 QY 1184 VCMAPRSSQSSSEGSF-----SQVSTMALHIAQADAE-----DSPPSLQORHS 1225
 Db 976 FLRSDHPAILRQAPRPGIHSRSPIDTSVLYTAVQPNESDNDYIIPLPKPDVADDEG 1035
 QY 1226 LAARYNNVVSFPCCLARGAETRGSSRMKTFEFPMTPTTYKGSVDNQTDSGMVLASEEFE 1285
 Db 1036 -----LPEGSPSLASS-----TLNEVNTSSTI-----SCDSPLEL-QEPPQ 1070
 QY 1286 QIESRHR-----QESGFCKGP 1302
 Db 1071 QAEPEAQLEQPDQSG--CPGP 1089

Db 7 LPWLWLGIVSEADLVSSYSMTPTTISITEEHI INAKDTLITTCRQHPLSWSPGARW 66
 QY 67 APATGDK-----DSEDTGV---VRCEGTADAPYCKVLLLEHVANDTSGVVCYKIK 117
 Db 67 TPVGRRRWNSQPOORPVAGNPEDECEGTGTPYCKVILVLTESQANDGYHCYKID 126
 QY 118 ARIGTTAASVYVRFDEQFINKPDTLLVNRKDMAMPVCLVSPGLNVLRSQSSVLW 177
 Db 127 AKIEGTTAVSAYIFVRFDEQFINKPDTLLVNRKDMAMPVCLVSPGLNVLRSQSSVLW 186
 QY 178 PDGEVWMDRRGLVPLTHDALYLCQETTWGDDQFVSNPFLVHITGNELYDQLPLR 237
 Db 187 PDKTIWONKKGQVPTQLIRSLFVQCETVINDKVFKNFIIHAGIELYDQLPK 246
 QY 238 KSELLVGEKLVNCTVWAEFNSGVTFDWDYPGKQAEGRKWPVRSQOOTHTELSILTI 297
 Db 247 KAMELLVGEKLVNCTVWAEFNSGVTFDWDYPGKQAEGRKWPVRSQOOTHTELSILTI 306
 QY 298 HNVSOHDLGYSVCKANNQIRFESTEVIVHNPFIISVWKLGPILVATAGDELKLPVK 357
 Db 307 HNVSOHDLGYSVCKANNQIRFESTEVIVHNPFIISVWKLGPILVATAGDELKLPVK 366
 QY 358 LAAYPPPEQWYKDGKALSGRSHALVKEVTEASTGTVTLALWNSAAGLRNLSLELV 417
 Db 367 VVAYPOPOFOMYKAGK-LIPKOSOSSMOIKDVAEHAGTITVLRNRLVGLERISLOLI 425
 QY 418 VNVPPQIHEKEASPSIYSRHSQALCTAYGVPLPLSIOWHWRPWPCKMFAQRSR-- 475
 Db 426 VNVPPRIHEKETSFSIYSRHSQALCTAYGVPLPLSIOWHWRPWPCKMFAQRSR-- 485
 QY 476 ----RQOQDLMPQCRDWRVNTQDANVPDIESLDTWTFEVEGKNTKSKLVIONANVSAM 531
 Db 486 HRAARRHORDRPECKMDKVSQDANVPDIESLDTWTFEVEGKNTKSKLVIONANVSAM 545
 QY 532 YKCVSNKYGDERLTYFYVTTIPDGFTEKSPSEELLEGQPVLLSQADSKYHELRWY 591
 Db 546 YKCIASNKVGRDERLTYFYVTTIPDGFTEKSPSEELLEGQPVLLSQADSKYHELRWY 605
 QY 592 RLNLSTLHDAGNPLLLDCKNVHLFATPLAASLEVAFCARHATLSLSIPVAREHGHY 651
 Db 606 RLNLSTLHDAGNPLLLDCKNVHLFATPLAASLEVAFCARHATLSLSIPVAREHGHY 664
 QY 652 VCEVQDRRSHDKCHKYLSVQALEAPRLTOLLVNVSDLEMOCLVAGAHAPSIW 711
 Db 665 VCEVQDRRSHDKCHKYLSVQALEAPRLTOLLVNVSDLEMOCLVAGAHAPSIW 724
 QY 712 YKDERLEESGVDLADSNOKLSIQVRREDAGRYLCSVCNKGCVNSASVAVGSEDK 771
 Db 725 YKDEKLVEEVSIGDLADFNORLSIQVRREDAGRYLCSVCNKGCVNSASVAVGSEDK 784
 QY 772 GSWIEVLTGTVAVFVFWLLLIIFCNRRPAHADIKTYLSIIMDPGEVPLEEQEYL 831
 Db 785 TNWEIVLTGTVAVFVFWLLLIIFCNRRPAHADIKTYLSIIMDPGEVPLEEQEYL 844
 QY 832 SYDASOWEFPRLHLGRVLGYGAFKGVVEASAFGHHKSSCDTVAVMLKEGATASERH 891
 Db 845 PYDSSKWEFPRLHLGRVLGYGAFKGVVEASAFGHHKSSCDTVAVMLKEGATASERH 904
 QY 892 ALMSELKILHIGNHLNVNLLGACTKPGPLMVIVFECKYGNLSNFKRAKDAFSPCAE 951
 Db 905 ALMSELKILHIGNHLNVNLLGACTKPGPLMVIVFECKYGNLSNFKRAKDAFSPCAE 964
 QY 952 KSPQRGRFRAMVELARLDR--RPGSSDRVLFAFARFKTEGGARRASPDQEAEDLMLSPUT 1010
 Db 965 KSPRLRQVQISIVEAVRADRRSRGTSDSAIFNRLMHK--SQTVPQI0EYDOLWQSPUT 1022
 QY 1011 MEDLVCSFQVARGMEFLASKCHRDOLAARNLLSFDVVKICDGLARDIYKDPDYR 1070
 Db 1023 MEDLVCSFQVARGMEFLASKCHRDOLAARNLLSFDVVKICDGLARDIYKDPDYR 1082
 QY 1071 KGSARLPKKNWAPESIFDKVYTTQSDVMSFGLVLEIFSLGASPYPGVQINEEFCORLKD 1130
 Db 1083 KGSARLPKKNWAPESIFDKVYTTQSDVMSFGLVLEIFSLGASPYPGVQINEEFCORLKD 1142

QY 1131 GTRRAPELATPAIRIMLNCWSDGPKARPAFSELVEIILGDLQGLQEEVEVCAPRS 1190
 Db 1143 GTRRAPEYTTABEYIRMLSCWHGDKPRTFTSDVLEILGNLLQ-ENVQEGKDRIPLND 1201
 QY 1191 SOSSEGSFQVSTALHIAQDAEDSPSLQHSLSLAARYNNVSPGCLARGAETRGSS 1250
 Db 1202 SHSEDDGFSQVSSAQO--NSDEEDFDMRIRCHLSLAARYNCVSPGCLTGMQIRCSS 1259
 QY 1251 RMKTFEFPPTPTTYKGSVDNODTSGWLASSEFEQIESHRHROESGFCCKGQCONVATR 1310
 Db 1260 RIKTFEFPHTMYKAHPDNQDTSNGLASEFEFEIENHRKEGFGSKGNRTAELSA 1319
 QY 1311 AHPSQRRRRPREGAR-GGOVFNSEYGESESPSEDEHSCPSA-----RVTFETD 1360
 Db 1320 EQSLRG-RCRPSYGOVGQGTTFYNSEYGESESEDRSCTPPAEGASPPALHASFSE 1377
 RESULT 2
 Q9PTLO PRELIMINARY; PRT; 1173 AA.
 AC Q9PTLO;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE FLK1 PROTEIN (FRAGMENT).
 OS Brachydanio rerio (zebrafish) (zebra dario).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Rasbora; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP MEDLINE=98294174; PubMed=9630750;
 RA Thompson M.A., Ransom D.G., Pratt S.J., MacLennan H., Kieran M.W.,
 RA Detrich H.W.I.I., Vall B., Huber T.L., Paw B., Brownlie A.J.,
 RA Oates A.C., Fritz A., Gates M.A., Amores A., Bahary N., Talbot W.S.,
 RA Her H., Beier D.R., Postlethwait J.H., Zon L.I.;
 RT "The cloche and spadetail genes differentially affect hematopoiesis
 RT and vasculogenesis.";
 RL Dev. Biol. 197;248-269(1998).
 RN [2]
 SEQUENCE FROM N.A.
 RA Thompson M.A., Ransom D.G., Pratt S.J., MacLennan H., Kieran M.W.,
 RA Detrich H.W.I.I., Vall B., Huber T.L., Paw B., Brownlie A.J.,
 RA Oates A.C., Fritz A., Gates M.A., Amores A., Bahary N., Talbot W.S.,
 RA Her H., Beier D.R., Postlethwait J.H., Zon L.I.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF180354; AAF03237.1;
 DR HSP: P08631; IAD5.
 DR INTERPRO: IPR000719;
 DR INTERPRO: IPR001245;
 DR INTERPRO: IPR001824;
 DR INTERPRO: IPR003006;
 DR PFAM: PF00047; Ig; 5.
 DR PFAM: PF00069; pkinase; 2.
 DR PRINTS: PR00109; TYRKINASE.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 FT NON_TER 1
 SQ SEQUENCE 1173 AA; 132446 MW; A0711FC3BC7583AA CRC64;

Query Match 29.7%; Score 2144; DB 13; Length 1173;
 Best Local Similarity 41.6%; Pred. No. 8.2e-170;
 Matches 502; Conservative 187; Mismatches 373; Indels 144; Gaps 35;
 QY 143 PDLTVNRKDMW-VPLCVISIPGLNVLTRSGSSVLPDQGE-VVWDRRGMVSTPL-H 199
 Db 4 PHLFAYENDPFFVPCRTTYPNQNVILETQMNPMDADDVGRGVQMDPKKGTIV--PLKPY 61


```

[2]
RP SEQUENCE FROM N.A.
RA Thompson M.A., Ransom D.G., Pratt S.J., MacLennan H., Kieran M.W.,
RA Detrich H.W. III, Vail B., Huber T.L., Paw B., Brownlie A.J.,
RA Oates A.C., Fritz A., Gates M.A., Amores A., Bahary N., Talbot W.S.,
RA Her H., Beier D.R., Postlethwait J.H., Zon L.I.,
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF178759; AAD56011.2; -
DR HSP: P06213; IIRK.
DR INTERPRO: IPR000719; -.
DR PRINTS: PF00069; pkinase.1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
FT NON_TER 1
SQ SEQUENCE 404 AA; 45281 MW; A149B836219EE4A9 CRC64;

Query Match 15.6%; Score 1128; DB 13; Length 404;
Best Local Similarity 54.8%; Pred. No. 1.le-85;
Matches 241; Conservative 43; Mismatches 96; Indels 60; Gaps 8;

QY 933 GNLSNFRKRAKDAFSPCAKSPQGRFRAMVEL--ARLDRRRRPGSSDRVLFAFRTKEG 990
Db 1 GNLSNFRKRAKREFFLPYRDRSPKTSQVRRMIEAGQASQSEHQPSTSS----- 48

QY 991 GARRASPDQAEADLWLSPLTWEDLVCFQVARGMEFLASRKCIRHDLAARNILLESVD 1050
Db 49 ---TNPRTVTDLMKPTLTIEDLVCFQVARGMEFLASRKCIRHDLAARNILSEN 105

QY 1051 VKICDGLARDIYKDPYVRKGSARLPLKMWAPESIFDKVYTTQSDVNSFGVLLWIEFSL 1110
Db 106 VKICDGLARDIYKDPYVRKGNALPLKMWAPESIFDKVYTTQSDVNSFGVLLWIEFSL 165

QY 1111 GASPYGVQINEEFCQRLDRGTRMRAPELATPAIRRMILNCWSDGPKRAPFSELVEILG 1170
Db 166 GASPYGQIODEFCRLKLDGTRMRAPDNASPEIYGINLACWQGEPRPTFPALVEILG 225

QY 1171 DLLOGRGLQEEVEVCMAPRSSOSSEGSFQSVQTMALHIAQADADSPS-----LQRHS 1225
Db 226 DLQENSLPE-----IPFNVSOSSEDDGFSQASS-----RPPSEIEFLACNT 268

QY 1226 LAARYNYNVSFFGCLARGAETRGSRMKTFFEFMTPTTYKGSVDNQDSDGHWLASSEFE 1285
Db 269 LPTRYNYNVPFAGCVNMGVSSSTCHSRVKTFFELPMWTSKTHQSDSDGHWLASDELE 328

QY 1286 QIESRHRSQSGFSCGKPGQNVAVTRAHPD-----SQRRRRPERGAR-GGQVF 1332
Db 329 RFEKHR-----GAMLTATTGQSTDRILSCPSVSSSGGGILLRPVFTQLSGQTF 380

QY 1333 YNSEYGEIPEPSEDEHCSPS 1352
Db 381 YNNEYGLSEEGVSGDYFSSS 400

RESULT 6
Q9VLQ8 PRELIMINARY; PRT; 1500 AA.
ID Q9VLQ8
AC Q9VLQ8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CG8222 PROTEIN.
GN CG8222.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;

Query Match 15.4%; Score 1108; DB 5; Length 1500;
Best Local Similarity 25.3%; Pred. No. 4.9e-83;
Matches 401; Conservative 221; Mismatches 563; Indels 398; Gaps 59;

QY 30 PTLNITESHVDTGDSLSISRCRGHPLEWAPCAQEPATGDKDSEDTGVVRCEGTD 89
Db 47 PLMTPCKSAIILDAQTSTLLCDEDEPMSWMTSQVYHVKSFTDTE-----PA 96

QY 90 RPYCKVLLHVEVHNDTGSYVCY----YKYI-----KARIEGTT---AASSTVFVRDF 135
Db 97 RPFCTSLHIEVTADYVAAYVCVTKSFQJAKEEQSDAMIELVNGQYASSIVYVND- 155

QY 136 EOPFINKPDTLLVNRKDAW-----VPCLVIPGLNLTIRSQSSVLPDQGVVWD- 186
Db 156 -----PDTLVLDSHNVYVTAQYTDVVPKAMPDTEVLLETSSNG-----ESTWKI 201

QY 187 DRGMLVSTPLLDAL-----YLCETTWGDDQDFLSNPLFVHITGNE 228
Db 202 SSKGOIQGNPKFEFDSVRYHPRGFTFRVDCVSGLYCKT-----LDYELIIDVTYPE 254
```



```
Db 2 RGA---RAWDFLVQLLRLVQTGSSQSPVSPSELSHPKSELIVSAGDEIRLFC 58
QY 52 RGHPLFNAWPGOAPATGDRSDTGVRDCEGTDAAPYCKVLLHEVHANDTGSVVC 111
Db 59 TDGSGVKWTF-----ETLGL--SENTHAE-----WIVEKAEAMTGNVTC 97
QY 112 YKYIKARIEGTAASSVVFVRDFEQPIKRPDITLLVNRKDMWVPCLVSPGL-VNLT 170
Db 98 -----TNEGGLSSIVFVRDPEKFLVDPPLVYKEDNDAL-VRCPLTDPVNTYSL 149
QY 171 SQSSVLPDQGVWDDRRGLV-STPLLDALYLQCTTGWODFLSNPELVHITGNEL 229
Db 150 GCEGRPLKDLTFVADPRAGITIRNVKREYHRLCHCSANOGGKSVLSKKTLLVRA--- 206
QY 230 YDIQLP-----RKSELLVGEKLVNCTVNAEFNSGVTFDH-----DYPGQAKRGK-- 278
Db 207 -AIRAVPVVAVKASYLLREGEFAVMCLI-KDVSSVSDSMWIRENSQTKAQVRKNSWHQ 264
QY 279 -----VPERSSOOTHTELSILTIRNVSOHDIGSVYCKANNGIQRFRETEVIVHENPFI 333
Db 265 GFENLRQR-----ITISARVNDSGVFMCMYANNTFGSANVTITLEVVDKGF 313
QY 334 SVEWLAGPILATA---GDELVKPLVLAAYPPPEF-QW-----YKDGKALS 376
Db 314 NI-----FPMNNTVFVNDGEDVDLIVEYAPKPEHROWIYMNRTATDKWEDYPKSENE 369
QY 377 GRHSPHALVKEVTEASTGYTILALWNSAAGLRNLSLELVNVVPPQIHEKEASSPSIYS 436
Db 370 NIRYVSELHLTKLGTGGTYTFLVSN--ADVNSVTFVNYNTKPEI-----LTHD 419
QY 437 RHSRQALCTATGYVPLPLSIQIWHRRWPTCKMFAQRSLRRQOQDLPOCRDMRAVTTQD 496
Db 420 RLNGHMLQCVAGFPEP-TIDWYFCGT-----EQRSVP----- 453
QY 497 AVNPISLDTWTEFVEGKNKTVSKLVIONANSAMYKCVSNKVGODERLIYFYVTTIPD 556
Db 454 -VGVPD-----VQIQNSSVSPFGKLYI----- 474
QY 557 GFTIESKPEELLEGOVLLSQADSYKVEHLRWYRLNLSLTHDAHGNPLLLDCKNVHLF 616
Db 475 ----- 474
QY 617 ATPLAASLEAVAPGARHATLSIPRAVEHEGYVCEVQDRSHDKHKKYLSVQALE 676
Db 475 -----HSSIDYS---AFKHNGTVEC---RAYNDVGKSSAFFEAKFE 510
QY 677 APRLTQNLTLVNVSDSLQCLVAGAHAPSIVWTKDERLLEKSGVDLADSNOKLSIQ 736
Db 511 QIHAHTLFTPLLTGFVIAAGMMCIIV-----MILTYK----- 542
QY 737 RVREEDAGRYLCVCAKGVNASSASVAVEGSDKSMETIVLGVTVAVFFVLLLI 796
Db 543 -----YL-----QKPMYEQV-----WKVVEI 559
QY 797 FCNRRPAHADIKTGYLSIIMDPGEVPLEDQCEYLSYDASQWEPFRERLHGRVLGYGAF 856
Db 560 -----NGNNYVIDTQLP-----YD-HKWEFPNRNLSFKCTLGGAGAF 596
QY 857 GKVEASAFGIHKGSSCDTVAVKMLKEGATASEHRAIMSELKLIHIGHLNVNLLGAC 916
Db 597 GKVEATATGLISDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHMINVLLGAC 656
QY 917 TKPGPLMIVTEPKYGNLSNFRKARDAF-----SPCAKSP 954
Db 657 T-IGGPLVITECCYGOLLNFLRRKDSFICSKQEDHAEALYKLLHKSKESSCSDSN 715
QY 955 EQGRFRAMVELARLDRRRPGSSDRVLFAFSTEGEGARRAS-----PDQAE 1002
Db 716 EY-----MD-MKPGSVYV-----PTKADKRARSAGISYIERDVTPTAMEDD 756
QY 1003 DLWLSPLTMEDLYCYFOVARGMEFLASRKCIRDLAARNILLSESDVVKICDFGLARDI 1062
Db 757 EL-----ALDLEDLLSFSQVAKAGMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDI 813
```

```
QY 1063 YKDDPYVRKSARLPLKWMAPESIFDKVYTTQSDVWSEFVLLWEIFSLGASPYGVOINE 1122
Db 814 KNDSSNYVGNARLPLVKWAPESIFNCVYTFESDVWSEYIGFLWELFSLGSSPYGMPVDS 873
QY 1123 EFCORLSDGTFRMAPELATPAIRRLMNCWSDGPKARPAFSELVEIL 1169
Db 874 KFYKIREGFRMUSPEHAPAEYMDIMKTCWDADPLKRPTEKQIVQLI 920
```

RESULT 8

```
ID Q9TQ01 PRELIMINARY; PRT; 964 AA.
AC Q9TQ01;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
GN KIT1*0201.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833(1998).
DR EMBL; AJ233229; CAA11197.1; -.
DR HSSP; P06213; IIRK.
DR INTERPRO; IPR00719; -.
DR INTERPRO; IPR001245; -.
DR INTERPRO; IPR001824; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; Ig; 2.
DR PRINTS; PF00069; pkinase; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Receptor.
FT NON_TER
SQ SEQUENCE 964 AA; 108315 MW; 996C3C46201358A8 CRC64;
```

```
Query Match 14.6%; Score 1055; DB 6; Length 964;
Best Local Similarity 27.0%; Pred. No. 6.2e-79;
Matches 337; Conservative 159; Mismatches 343; Indels 408; Gaps 42;
```

```
QY 3 RGAALCLRLWLCGLLGDGL-----VSGYSMTPTPTNITEESVITGDSLSISC 51
Db 2 RGAAR---RAWDFLVQLLRLVQTGSSQSPVSPSELSHPKSELIVSAGDEIRLFC 58
QY 52 RGHPLFNAWPGOAPATGDRSDTGVRDCEGTDAAPYCKVLLHEVHANDTGSVVC 111
Db 59 TDGSGVKWTF-----ETLGL--SENTHAE-----WIVEKAEAMTGNVTC 97
QY 112 YKYIKARIEGTAASSVVFVRDFEQPIKRPDITLLVNRKDMWVPCLVSPGL-VNLT 170
Db 98 -----TNEGGLSSIVFVRDPEKFLVDPPLVYKEDNDAL-VRCPLTDPVNTYSL 149
QY 171 SQSSVLPDQGVWDDRRGLV-STPLLDALYLQCTTGWODFLSNPELVHITGNEL 229
```

```
Db 150 GCGKPLPKDITVADPKAGIIRNKKREYHRLCLHCSANOGKSVLSKKFTLKVPA--- 206
QY 230 YDQLLP-----RKSELLVGEKLVNCTVWAEFNSGVTFDM-----DYPGKAERGM-- 278
Db 207 -AIRAVPVAVASKASYLLREGEFAVCLLI-KDVSSVSDSMWIRENSQTKAOKVRNSHQ 264
QY 279 -----VPERSOQTHLSILLIHNVSQHDLSGYCKKANGIORRESTEIVHNPPI 333
Db 265 GDFNLRQER-----LTISSARVNDSGVFYCIANNTPGSAVITITLEVVDKGI 313
QY 334 SVEMLAGPILDEATA---GDELVKLPVLAAYPPPEF-QW-----YKDGKALS 376
Db 314 NI-----FPMNTTFVNDGEDVDLIVEYEAYPEKPEHQMIYMNRTATDKWEDPKSINES 369
QY 377 GRISPHALVKEVTEASTGYTITLALNWSAAGLRNLSLELVNVPQIHEKEASSPSIYS 436
Db 370 NIRYVSELHLRLKGTGGTYTFLVSN--ADVNSVTFVFNVTNKPFI-----LTHD 419
QY 437 RHRQALCTCTAYGVPLPLSLQIWHRRPWTCKMEAFQSLRRRQOQDLMPQCRDWRVTTOD 496
Db 420 RLNGMLQCVAAAGPEP-TIDWYFCPT-----EQRCVSP----- 453
QY 497 AVNPESLDTWTVEVEKNTVSKLVTONANVSAMKCVVSNKVGQDERLIYFYVTIPD 556
Db 454 -VGPVD-----VQIONSVPSPFGKLV----- 474
QY 557 GFTTESKPSBELGQPVLSSQADSYKYEHLRWYRLNLSTLHDAHGNPLLLDCKKNVHLF 616
Db 475 ----- 474
QY 617 ATPLAASLEVAPOARHATLSLTPRVAPEHGYHVEQVDRSHDKHKKYLSVQALE 676
Db 475 -----HSSIDYS-----AFKHNGTVEC-----RAYNDGKSSAFNEAFKE 510
QY 677 APRLTQNLTLVNVSDSLEMOCLVAGAHAPSIVYKDERLLEKSGVDLADSNOKLSIQ 736
Db 511 QIHAHTFTPLIGFVAGNWCIIV-----MILTYK----- 542
QY 737 RVREEDAGRYLCSVCNAKGCNVSSASVAVGSEDKGSNEIVILVGTGVIAVFFWVLLLI 796
Db 543 -----YL-----QKPMYEQV-----WKVVEI 559
QY 797 FCNMRPAHADIKTYLSIIMDGPVPLEEOCEVLSYSDASOWEPFRERHLGRVLGYGAF 856
Db 560 -----NGNNVYIDPTQLP-----YD-HKWEPPRNLSEFGTLGAGAF 596
QY 857 GKVEASAFGIHKSSCDTAVAKMLKEGATASEHRLMSLKIILIHGNHLNVNVLGAC 916
Db 597 GKVEATAYGLIKSDAAMTAVAKMLKPSAHLTEREALMSLKVLSYLGNNHNIIVNLGAC 656
QY 917 TKPOGPLMIVVEFCYKGNLSFLAKRDFA-----SPCAEKP 954
Db 657 T-IGPTLVITECYGDLNLFRRKRDSTFCQKQEDHAEALYKNLLHKSCESSDSTN 715
QY 955 EQRGRFRAHVELARLDRRPGSSDRVLFAFPFKTEGGARRAS-----PQGEAE 1002
Db 716 EY-----MD-MKPGSVYV-----PTKADRRRSARIGSYTERDVTIPAIEDD 756
QY 1003 DLWLSPLTHEDLVCSFQVARGMEFLASRKCIHRDLAARNILLSESDVVKICDGLARDI 1062
Db 757 EL---ALDELLGSFYQVAGHAFASKNCIHRDLAARNILLTHGRITKICDGLARDI 813
QY 1063 YKDPDYVRKGSARLPKKNWAPESIFDKVYTTQSDVSWSGVLLWEIFSLGASPGVQVINE 1122
Db 814 KNDNSYVVKGNARLPVKWNPAPESIFNCVYTFESDWSVSYGIFLWELFSLGSSPYGMPVDS 873
QY 1123 EFCQRLRGCTWRAPELATPAIRIMLNCWSDPKAPAFSELVEIL 1169
Db 874 KEYRKIEGFRMLSPHAPVEMIDIMKTCWDADPLKRPFTFKQIVOLI 920
```

RESULT 9

Q9TQO0

```
ID Q9TQO0 PRELIMINARY; PRT; 964 AA.
AC Q9TQO0;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE MAST/STEM CELL GROWTH FACTOR RECEPTOR (FRAGMENT).
GN KIT1*0202.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;
RA Marklund S., Kiljas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT *Molecular evolution of the dominant white phenotype in the domestic
RT pig.;
RL Genome Res. 8:826-833(1998).
DR EMBL; AJ223230; CAA11198.1; -.
DR HSSP; P06213; IIRK.
DR INTERPRO; IPR000719; -.
DR INTERPRO; IPR001245; -.
DR INTERPRO; IPR001824; -.
DR INTERPRO; IPR003006; -.
DR PFAM; PF00047; Ig; 2.
DR PFAM; PF00069; pkinase; 2.
DR PRINTS; P000109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Receptor.
FT NON_TER 964
SQ SEQUENCE 964 AA; 108287 MW; 1B21292A962E9191 CRC64;
```

Query Match 14.6%; Score 1053; DB 6; Length 964;

Best Local Similarity 27.1%; Pred No. 9.2e-79;

41;

Matches 336; Conservative 158; Mismatches 352; Indels 394; Gaps

```
QY 3 RGAALCLRLWLCGLDGL-----VSGYSMTPTPLNITEESHVIDTGDLSLISC 51
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 RGAR---RAWDFLVQLLLRVQTGSSQSPSVSPEELSPPSIHPAKSELIVSAGDEIRLFC 58
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 52 RGQHPLEWAPGQAEPATGDKDSEDTVVVRCEGTDRPYCKVLLHVEVHANDTGSVVC 111
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 59 TDPGSKVTF-----ETLQGL--SENTHAE-----WIVEKAEAMNTGNITC 97
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 112 YYYIKARIEGTTAASVYVVRDFEQPFINKPDTLLVNRKDAWVPCVLSIFCL-NVTILR 170
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 -----TNEGGLSSIIYVVRDPEKFLVDPPLYGKEDNDAL-VRCPLTDPVNTSLT 149
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 171 SQSSVLWPGQGVVWDDRRGLMV-STPLLDALYLOCETTWGDDFLSNPFLVHIIGNEL 229
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 GCGKPLPKDITVADPKAGIIRNKKREYHRLCLHCSANOGKSVLSKKFTLKVPA--- 206
QY 230 YDQLLP-----RKSELLVGEKLVNCTVWAEFNSGVTFDM-----DYPGKAERGM 280
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 -AIRAVPVAVASKASYLLREGEFAVCLLI-KDVSSVSDSMWIRENSQTKAOKVRNSW-- 262
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 ERSQOQTHLSILLIHNVSQHDLSGYCKKANGIORRESTEIVHNPISVEMWLG 340
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 --HOGDFNLRQEKLTISSARVNDSGVFYCIANNTPGSAVITITLEVVDKGIINI---P 316
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 PILEATA---GDELVKLPVLAAYPPPEF-QW-----YKDGKALSGRHSHPA 383
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

Db 474 ----- 473
QY 698 QCLVAGAHAPSIWYKDERLLEKSGVDLADSNOKLSIQRVREEDAGRYLSCVNAKGC 757
Db 474 -----GA-----VEVESFTVGLSHRMTVE-----CVAENLVGV- 503
QY 758 NSSASVAVGSEKSGMEIVILVG-TGVIAVFFWVLLILLIFCNRRPAH-----ADIKT 810
Db 504 -SSDTFTVEVSD---KLFSTLTIGAGVLAIFLLLVFLLYKOKPRFEIRWKIIARE 559
QY 811 GYLSIIMDGEVPLEEQCEYLSYDASQWEEFRERLHLGRVLGYGAFGVVSEASAFGIHK 870
Db 560 GNNYTFIDTQLPYNE-----KWEFPDKLKLKVLGAGAFGVVSEATFGL--G 607
QY 871 SSCDT---VAVKMLKEGATASEHRALMSSELKILIHGHNLVNVNLLGACTKPOGLVIV 927
Db 608 EDKDNTRLVAVKMLKANAHSDEREALMSSELKILSHLGHQNTVNLGACTY--GGPVLVIT 666
QY 928 EFCYGNLSNLFRAKRAFAFSPCAEKSP-----QGRFRAMVELARLD--- 970
Db 667 EYCSLGDLLNLFLOKAEFTVNLVMIPEIMENSNDYKNICNKKWYIRSDSGISSTSSSTY 726
QY 971 -RRPGSSDRVLFAFRTKTEGGARRASPOEAEDLWLSPLTLMEDLVCSYFQVARGHEFLA 1029
Db 727 LEMRPSQOQSHI-----EASGRKSLCEDNGD--W--PLDIDLLRFSLQVAOGLDFLA 774
QY 1030 SRKCIHRDLAARNILLSESDVYKICDFGLARDIYKDPDVVRKGSARLPLKWNAPESIFDK 1089
Db 775 SRKCIHRDVAARNVLLTDKRAKICDFGLARDIMDSNVYVGNARLPLKWNAPESIFDC 834
QY 1090 VYTTQSDVMSFGVLLWEIFSLGASYPVGVQVINEFCQRLRDGTRMRAPELATPATIRML 1149
Db 835 VYTVQSDVMSYGLLWEIFSLGASYPVGVQVINEFCQRLRDGTRMRAPELATPATIRML 894
QY 1150 NCWGGPKAPAPSELVELGLDGLQGRLOE-----EEVCMAP----- 1188
Db 895 MCWNLTEPTEFMSIQMNLNLLGGQDEKLIYRNVOPEQVAGEACEDEFKRYDPCE 954
QY 1189 RSSQSEE 1196
Db 955 RSCDHEE 962

RESULT 11
Q918N6 PRELIMINARY; PRT; 977 AA.
ID Q918N6
AC Q918N6
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE FMS.
GN FMS.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbora; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB;
RA Parichthys D.M., Paw B., Ransom D., Zon L.I., Johnson S.L.;
RT "An orthologue of the kit-related gene fms is required for development
RT of neural crest-derived xanthophores and a subpopulation of adult
RL melanocytes in the zebrafish, Danio rerio."
DR Development 127:0-0(2000).
EMBL: AF240639; AAF76872.1;
FT VARIANT 238 238 N -> K.
FT VARIANT 615 615 V -> M.
SQ SEQUENCE 977 AA; 110187 MW; C91A2F339E746A58 CRC64;

```

Query Match

14.6%; Score 1050; DB 13; Length 977;

```

Best Local Similarity 26.1%; Pred. No. 1.7e-78;
Matches 329; Conservative 178; Mismatches 356; Indels 398; Gaps 42;

QY 11 LMLCGLLGLGSLGYSMTPTLNLITE-----ESHVIDTGLSISCRGQHPLEWHPGAQ 65
Db 5 LFLIGILLGOVQGS--EPRIRLSGALAGTAVILGSGPLQVLGEGDGVTF----- 56
QY 66 EAPATGDKD---SEDITGVVRDCEGTDAAPYCKVLLLHEVANDTGSVYCVYKIKARIEG 122
Db 57 -LPLAKHKYISKEVGKIRS-----FRVKITVDTGYIKCYMNGN---DS 100
QY 123 TTAASSYVYFDEFEQPIKRPDILLVNRKDA--MWVPLVSIIP-GLNLTURSSSVLWPD 179
Db 101 NLSSSVHVFVRDVRVLPSPSTSLRYVVRKEGEDLLPCLLTDEATDFTFMDNGSAAPY 160
QY 180 GOEVDWDRGMLVSTPLH--DALVLOCETTWGDDFLSNFLVHITCNELVDIO--- 233
Db 161 GNNITYDPKGVLLRN--VHPGFNADYICARI--GGAEKVKIFSIINI-----IQRLR 210
QY 234 -----LLPRKSLELLVGEKLVLCNTV--WAEFNSGVTFDMYDPCQAKERGKVPERRSQOT 287
Db 211 FPPVYVYKRYVYKLVGERLQISCTTNNPNFYVNT--WTHSSRM-----LPRAEKST 262
QY 288 -----HTLSLTLTIHNYSOHDLGSYVCKANNGIORFRESTEIVHENPESIVE----- 336
Db 263 MEGRLAIESILTIIPSVOLSHGTGNTCTGQNEAGANSSTTQLLVVEEPIKLSPLSKL 322
QY 337 WLKGPILAEATAGDELKLVPLKLAAYPP-----PEFQWKDGKALSGRHS 380
Db 323 THRGSLIEVSEGD--VDLGVLEIAYPLTSHKWTPTSHNASLPENRFNHNDRYEA--- 378
QY 381 PHALVLEVEATAGTGTTLALWNSAAGLRNLSLELVVNPVPPQIHEKEASPSIYSHSR 440
Db 379 --LLLRLELFEETGQVTLNYSKMS--ASITFDIKMTKPVARVWENTT----- 427
QY 441 QALTCTAYGVPLPSIOHWRPMTPCKMFQSRRLRRQQDMLPQCDWRVAVTTQDA--V 498
Db 428 --LSCRSYGYPAP--SILWY-----OCTGIRTCENTTDL 459
QY 499 NPISLDTWTFEVEGKNT-----VSKLVTONANVSAMKYCVVSNKVKQDERLIYFVVT 553
Db 460 QP1Q-----TOTVEFQKESFGAVGESVLTVPNRRMTVVCVAFNLYGQGS----- 505
QY 554 IPDGFTEISKPESELLEGQVLLSCQADSYKYEHLRWYRLMLSTLHDAHGNPLILDCKNV 613
Db 506 --DTFSME----- 511
QY 614 HLFATPLAASLEEVAPGARHATLSLIPRVAPEHEGHVCEVQDRRSHDKCHKYLSVQ 673
Db 512 ----- 511
QY 674 ALEAPRLTQNLTDLLVNVSDSLEMOCLVAGAHAPSIWYKDERLLEKSGVDLADSNOKL 733
Db 512 -----VSDQI----- 516
QY 734 SIQVRVEDAGRYLSCVNAKGCNVSSASVAVGESDKGMEIVLVTGTVIAVFFWVLL 793
Db 517 -----FTSAMC-----GSTVAVVVLGL-----LLI 536
QY 794 LLIFCNMRPAH-----ADIKTGYLSIIMDGEVPLEEQCEYLSYDASQWEEFRERLHL 847
Db 537 FMIYKQKPRFEIRWKIIETATNGNNTYFIDTQLPYNE-----KWEFPDKLKL 586
QY 848 GRVLGYGAFGVVSEASAFGIHKSGCDTAVVAVKMLKEGATASEHRALMSSELKILIHGHNL 907
Db 587 GKTILGAGAFGVVSEATVGLGREDNITRVAVKMLKASAHDPDEREALMSSELKILSHGQK 646
QY 908 NVVNLGACTKPOGLVIVLVECKYGNLSNLFRAKRD---AFSPCAEKSPQGRGFRAMV 964
Db 647 NIVNLLGACTH--GGPVLVITEYCCCHGDLNLFKRAEFLNFMVITPFPEDPTDKNVS 705
QY 965 ELARLDRRPG-----SSDRVLFAR--FSKTEGGARRASPOEAEDLWLSPLTLMEDLVCSY 1018

```

Db 706 TERMFVRSDSGISCTSDHVDYLDMPVTSRPTNSALDSSSEQ-EDSW--PLDMKDLRLFS 762
QY 1019 FOVARGMEFLASRKCIRDLAARNILLSSESVVVKICDFGLARDIYKDPDYVRKGSARLPL 1078
Db 763 SOVAGGLDFLAKKCIHRDVAARNVLLTNSRVAKICDFGLARDIMDSNTVYVGNARLPV 822
QY 1079 KWMAPEISFDKVTYTQSDVWSFGVLLWEIFSLGASPGVGVINBEFCQRLDRGTRRAPE 1138
Db 823 KWMAPEISFEVYTVQSDVWSYGLWEIFSLGASPGVGNILVDSKFYKMIKCGVOMSRPD 882
QY 1139 LATPAIRRLMNCWGPDKAPPAESELVE---ILGDLLOGRGLQEEFEVCMAPRSSQSS 1194
Db 883 PAPPEMTIMKCNWLDAAERPTFSKISQMIQRMIGTSEQQDQEQYKNIPTAEAEQOL 942
QY 1195 E 1195
Db 943 E 943
RESULT 12
ID Q63116 PRELIMINARY; PRT: 978 AA.
AC Q63116;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE C-KIT RECEPTOR TYROSINE KINASE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID:10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE-92003944; PubMed-1912577;
RA Tsujimura T., Hirota S., Nomura S., Niwa Y., Yamazaki M., Tono T.,
RA Morii E., Kim H., Kondo K., Nishimune Y., Kitamura Y.;
RT "Characterization of Ws mutant allele of rats: a 12-base deletion in
RT tyrosine kinase domain of c-kit gene."
RL Blood 78:1942-1946(1991).
DR EMBL: D12524; BAA02094.1; .
DR HSP; P06213; IIR3.
DR INTERPRO: IPR000719; .
DR INTERPRO: IPR001245; .
DR INTERPRO: IPR001824; .
DR INTERPRO: IPR003006; .
DR PFAM: PF00047; ig; 2.
DR PFAM: PF00069; pkinase; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
SQ SEQUENCE 978 AA; 109341 MW; 0958C33F19889051 CRC64;

Query Match 14.5% Score 1044; DB 11; Length 978;

Best Local Similarity 26.2% Pred No. 5.3e-78;

Matches 328; Conservative 167; Mismatches 340; Indels 418; Gaps 42;

QY 3 RGA---ALCLRMLWCLGLDGLVSGYSM-----TPPTLNITEESHVIDTGSLSISCR 52
Db 5 RGAWDLLCVLLVLLRG-----QTGTSQPSASPGSPSPSIOPAQSELIVEAGDTIRLTCT 59
QY 53 GQHPLENAWCAQAPATGDKDSEDGTGVRCDEGTDPARYCK--VLLHVEHANDGTGSV 110
Db 60 DPAPVKKTF-----EFLDVRIENKSEWIREKAEATHTKYIT 96
QY 111 CYKYIKARIEGTTAASS-YVFRD-----FEQPFINKPDTLLVNRKDAWVPCLVISIP 163
Db 97 C-----VSGSLRRSSIVFVRDPAVLVLGLPLFGKED-----NDAL-VRCPLTDP 141
QY 164 GLNVLRSQSVLWPDQGVWD-----DRRGLV-STPLLDHALYLCQETTGWGDQDFL 216

Db 142 QV-----SNYSLIECDGKSLPTDLKFVPNPKAGITIKNVKRAYHRLCIRCAAQREGKMR 196
QY 217 SNPFLVHI-TGNELYDIQLLPRESLELLVEKLVNCTVWAEFNSGVTFDW-----DYP 269
Db 197 SDFTLKVRARAIKAIPIVSVPTSHLLKEGDTFTVICTI-KDVSTSVDSMMIKLNPQPS 255
QY 270 GKAERGKWPERRSQOHTHELSSILTIHNVSOHDGLSGSYVCKANRQIRPRESTEVIVHE 329
Db 256 KAQVKNSM---HQGFNVYERQETLTISSRVNDSGVFCYANNTFGSANVTITLKVYE 311
QY 330 NPFSISVENLKGPILEATAGDELVKPLAAAYPPPF-QWYKDGKALSGRHSIPA----- 383
Db 312 KGFIFPVKNTFTVFDG-ENYDLVFEAFEPKPEHQOQIYMNRTPTNRGEDIYKSDNQ 370
QY 384 -----LVLKVEVTEASTGTYTTLALWNSAAGLRNRNLSLELVNVVPPOLHEKEASSPSIY 435
Db 371 SNIRYVNELELRLKLGTEGTYTFLVNS--DVSASVTFDYVNTAPEI-----LTY 420
QY 436 SRHSQALCTCTAYGVLPLPSIOHMRPMTPCMKFAQSRSLRRROODLMPQCRDWRVTTQ 495
Db 421 DRLMNGRLQCVAAAGFPEP-TIDWYF-----CTCAEQR----- 451
QY 496 DAVNPIESLDTWTEFVEGKNKTVS---KLVIQANVSMYK-----CVVSNKVGQDERL 546
Db 452 -CTVPVPPVD-----VOIONASVPFGKLVQSSIDSSVFRHNGTVECKASNAVGSASF 505
QY 547 IYFVVTTPDGFTIESKPSLELGGQPVLLSCQADSYKYEHLNRYRLNLTSLHDAHGNTL 606
Db 506 F-----NPAFKNSKEQI---QP----- 520
QY 607 LDCKNVHLFATPLAASLEEVAPGARHATLSISIPRVAPEHGHYVCEVQDRRSHDKHCH 666
Db 521 -----HTLFTPLL----- 528
QY 667 KKYLSVQALEAPRLTQNLTLVNVSDSLEMQCLVAGAHAPSIVWKDERLLEEKSCVDL 726
Db 529 ----- 528
QY 727 ADSNOKLSIORVEEDAGRYLCVCNKAQGVNSASVAVEGSEDKGSMETVILVGTGVTA 786
Db 529 -----IGFVVTAGLWG 539
QY 787 VFWVLLLLIFCNMRPAH-----ADITGYLSITMDPGEVPLEBOCEVLTSDASQWEP 840
Db 540 I---IVNVLAYKYLQKPMYEVQNKVVEEINGNNVYIDPTLP-----YD-HKWEF 586
QY 841 PRERHLGRVLGYGAFGVVEASAFGIHKGSSCDTVAVKMLKEGATASBEHALMSKLTIL 900
Db 587 PRNLSFGKTLGAGAFGVVEATAYGLIKSDAANTVAVMLKPSAHLTEREALMSKLV 646
QY 901 IHIGNHLNVNVLGACTKPGPLMVIWEFCYKGNLSNFLRAKRDAF----- 946
Db 647 SYLGNHNMIVNLLGACT-VGGPTLVITEYCCYGDLELFLRRKRDSFISKQEOBDAALY 705
QY 947 ---SPCAEKSPQGRGRFAM-----VELARLDRRPGSSDRVLFAFPKTECGARRASP 997
Db 706 KNLHKSCKSDSSNEYMDMKPGVSVVPTKTKDRRSARIDSYI-----ERDVT 755
QY 998 D-OEAEEDLMLSPITMEDLVICYSFQVARGMEFLASRKCIRDLAARNILLSSESVVVKICDF 1056
Db 756 AIMEDDEL---ALDLEDLLSFSQVAKGNAFLASKNCIHRDLAARNILLTHGRITKICDF 812
QY 1057 GLARDIKDPDYVRKGSARLPLKWAPEISFDKVTYTQSDVWSFGVLLWEIFSLGASPTP 1116
Db 813 GLARDIRNSDNTVYVGNARLPVKNWAPESIFNCVYTFESDVWSYIGIFLWELFSLGSSPY 872
QY 1117 GVQINEEFQRLDRGRMRAPELATPAIRRLMNCWGPDKAPPAESELVEIL 1169
Db 873 GMPVDSKFYKMIKCGVOMSRPD-----DPLKRPVQVQLI 925

RESULT 13
Q63702

RL Genomics 39:216-226(1997).
DR EMBL; U63834; AAC50969.1; --
DR HSSP; P06213; LIRK.
DR INTERPRO; IPR000719; --
DR INTERPRO; IPR001245; --
DR INTERPRO; IPR001824; --
DR INTERPRO; IPR003006; --
DR PFAM; PF00047; 19; 2.
DR PFAM; PF00069; PKinase; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
SQ SEQUENCE 972 AA; 109450 MW; D59DEF9AF761FDA CRC64;

Query Match 14.4%; Score 1035.5; DB 4; Length 972;
Best Local Similarity 26.8%; Pred. No. 2.7e-77;
Matches 329; Conservative 158; Mismatches 369; Indels 371; Gaps 39;

QY 3 RGA--ALCLRLMLC---LGLLDGLVSGYSMTPTLTNITEESHVIDTGSLSISCRGQHPL 57
DB 5 RGAWDFCLVLLLLLRVQTGSSPSVSPGEPSPSIHPGKSDLIVRGDEIRLCTDPGFV 64
QY 58 ENAWPQAQAPATGDKDSEDGTGVRRDCEGTDAKYCKVLLLHEVANDTGSVYCYKIK 117
DB 65 KTFEILDET---NENKQEWITEKAEAT-----NTGRYTCNKH-- 101
QY 118 ARTEGTAASSYVDFEQPFINKPDTLLVNRKD-AMVPCLVSLPGL-NVTLRSQSSV 175
DB 102 -----GLSNSIYVDFRPAKLEU--VDRSLYKEDNDTLVRCPLDPEVTNLSLKGCGK 154
QY 176 LWDGQGVWDDRRGLV-STPLLDHALYLOCTETWGDQDFLSNPFLVH-TGNELYDIO 233
DB 155 PLPKDLRFIPDPRAGIMIKSVKRAYHRLCLHCSVDQEGKSVLSEKFIKVRPAFAVPV 214
QY 234 LIPKSLLELVGKLVNLTVAEFNSGVTFWDYFGKQAEKRGKWPERSQOOTH---T 289
DB 215 SVSKASYLLREGGEFTVCTI-KDVSSVYSTKRENSQTK-----LQEKINSHHGDFNY 269
QY 290 ELSISILTHNVSHOHLGYSVCKANNGIQRFRETEVIVHENFISVEMLKGPILENTA-- 347
DB 270 ERQATLTISARYNDGVPFCYANNFTGSAVNTTLEWDKGFINI---FPMINTTVEV 325
QY 348 -GDELVKLPVLAAYPPPEQ-----W-YKDGKALSGRSHALVLEKYTEA 392
DB 326 NDGENVDLIVEYFAFPKPEHQOYIYNNRTFTDKWEDYPKSESNIRYVSELHLTLKGT 395
QY 393 STGTYTTLALNSAAGLRNLTSLVNVNVPQIHEKEASSPSIYSHRSQALCTAYGVPL 452
DB 386 EGGTYTFLVNS--DVNAALAFNVYVTKPEI-----LTYDRLVNGMLQCVAAAGPE 435
QY 453 PLSTOWHWRPWPCKMFAQSLRRRQOOLMPQCRWRVTTQDANVPISLSTWTFEVE 512
DB 436 P-TIDWYFCPCTEQRCSA-----SVLPVD-----VQTLN 463
QY 513 GKNTVSKLVTONANVAMTK-----CVVSNKVGDERLIYFVTTIPDGFTIESKPSE 566
DB 464 SSGPFFKLVQSSIDSASAKHNGTVCKAYNDVGKTS--AYF----- 504
QY 567 ELLEGQPVLLSCQADSYKYEHLRWYRLNLTLDHAHNPDLCKKNVHLFATPLAASLEE 626
DB 505 -----NFAFKEQIH-----PHTLFTPLLIGFVI 527
QY 627 VAPGARHATLSLIPRAPEHEGHVYCEVDQDRSHDKHCKKYLVSQALEAPRLTONLTD 686
DB 528 VA----- 529
QY 687 LLVNVSDLEMQCLVAGAHAPSIVWYKDERLLEKSGVDLADSNOKLSIORVREEDAGRY 746
DB 530 -----GMCIIV-----MILTYK-----Y 543
QY 747 LCSVCNAKGCNVSSNASVAGESEDKSGMEIVLVGTGVIAVFFVLLILLIFCNMRRPAHA 806

DB 544 L-----OKPMYEVQ-----WKVVEEL----- 559
QY 807 DIKTGLSLIIMDPGEVPLEEOCEVLSYDASOWEFPFRERLHLGRVLGYGAFGVKEASAFG 866
DB 560 ---NGNYYVIDTQLP-----YD-HKWEFPNRLSFGKTLGAGAFGVKEATAYG 606
QY 867 IHKSSCDTVAVKMLKEGATASEHRLMSELKILIHGNHLNVNLLGACTKQGPPLMWI 926
DB 607 LIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHNMNIVNLGACT-IGGPTLVI 665
QY 927 VEFCKYGNLSNLFRAKRDAT-----SPCAEKSPEQGRFRAM- 963
DB 666 TEYCICYGDLNLFRRKKRDSFICSKQEDHAPALYKNLLHSKESCSDSSTNEYHDMRPGVS 725
QY 964 -VELARLDRRRPGSSDRVLFARFASKTEGGARRASPOQAEEDLWLSPLTMDLVCIYSQVA 1022
DB 726 YVFTKADKRR-----SVRIGSYIERDVTFAIMEDDELALDL-----EDLLSFYQVA 773
QY 1023 RGEFLASRCKIHRDLAARNILLSESDVKICDGLARDIYKOPDYVRKGSARLPLAWMA 1082
DB 774 KGMFLASKNCIHRDLAARNILLTHGRITKICDGLARDIKNDSNYVVGARLPLAWMA 833
QY 1083 PESIFDKVYTTQSDVWSFGVLLWEIFSLGASPYPGVOINEEFQORLDRGTMRAPELATP 1142
DB 834 PESIFNCVYTTESDWSYGFIFLWELFSLGSSPYPGMPVDSKFYKMIKEGFRMLSPHAPA 893
QY 1143 AIRIMLNCWSGDKAPAFSELVEIL 1169
DB 894 EMYDIMTKCDADPLKRPTEKQIVQLI 920

RESULT 15
O77589 PRELIMINARY; PRT: 945 AA.
AC O77589;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE TYROSINE KINASE RECEPTOR HOMOLOG (FRAGMENT).
GN KIT.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Lear T.L., Adams M.H., McDowell K.J., Ponce De Leon F.A.,
RA Sullivan N.D., Coogly L., Ferguson E., Chambers T.M., Bailey E.;
RT "Chromosomal location of genes for ESR, ETS2, KIT, MXI and PGR
in the horse, Equus caballus.";
RL Proc. Xth Amer. Coll. Dom. Anim. Cytogenet. Gene Mapp. 0:0-0(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;
RT "An Equine sequence homologous to the c-KIT-proto-oncogene mapped to
chromosome 3q13-21.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF055037; AAC23611.1;
DR HSSP; P06239; 3LCK.
DR INTERPRO; IPR000719; --
DR INTERPRO; IPR001245; --
DR INTERPRO; IPR001824; --
DR INTERPRO; IPR003006; --
DR PFAM; PF00047; 19; 2.
DR PFAM; PF00069; PKinase; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

FT NON_TER 945 945
SQ SEQUENCE 945 AA; 106142 MW; 7DB424A89DF5C340 CRC64;

Query Match 14.38; Score 1031.5; DB 6; Length 945;
Best Local Similarity 27.04; Pred. No. 5.5e-77;
Matches 333; Conservative 169; Mismatches 349; Indels 381; Gaps 43;

QY 3 RGA--ALCLRLWLC---LGLLDGLVSGYSMTPTPTLNTTESHVDTGDSLSISCRQHPL 57
DB 5 RGWDFLCVLLFRVQTGSSQSVSPGELSPPSIHPAKSELIVSGDEIRLUCADPGFV 64

QY 58 EWAPGAQAPATGDKDSEDTGVVRDCEGTDPYCKVLLLHEVHANDTGSVYCYKYIK 117
DB 65 KWTF-----ETWQLSNTH-----KENVTEKAEATNTGSYTC----- 97

QY 118 ARIEGTTAASYVVRDFEQPFINKPDTLLVNRKDAWVPCLYSIPGL-NVTLRSOSVYL 176
DB 98 -TNEGGLUSSSIYVVRDPAPKLFDFSLYKESSDTL-VRCPLTDPEVTNYSLMACEGKS 155

QY 177 WPDQOEYVMDRGMV-STPLLDALYLQCEITWGDODFLSNPFLVHITGNELYDIQLL 235
DB 156 LPKDLTFVADPKAGITIRNKREYHRLCLRSADKDGKSVLSNFKLVRA-----AIRAV 211

QY 236 P-----RKSLELVGKLVNCTVMAEFNSGVTFDWDYPGKQAEGRKWVPERSSQOHT- 289
DB 212 PVSVSKASYLLREGFEFSYTCIL-KDVSSSVDSMW-----IRENSRTEQVKSSSSHQ 265

QY 290 ELSI-----LTIHNVSHDLGVSVCKANNGIQRFRESTEVIHNPFIISVEMWLGPILEA 345
DB 266 DFNVRQERLITISPARVNDGVPFACVANNFTGSANVTTLLEVDKGCINV-----FPMNT 321

QY 346 TA-----GDELVKLPVLAAYPPPEFQ-----W--YKDGKALSGRSHSPHALVKE 388
DB 322 TVFVNDGENVDLIVESEYKPEHQWYMNRTSTDKMEDYPKSENESTIRYVSELRLTR 381

QY 389 VTEASTGYTLALWNSAAGLRNRISELVNVPPOIHERKASPSIYSHRSOALCTAY 448
DB 382 LKGTGEGTYTLVSNSS--DVDSSVTNVVYNTKPEILTR-----RLMNGMLQCVAA 431

QY 449 GVPLPLSIQHWHPWTPCKMFAORSRRRQOQDLMPQCRDWRVAVTTODAVNPTESLDTWT 508
DB 432 GFPEP-TIDWYFCPGT-----EQRCSIP-----VGPVD----- 458

QY 509 EFVEGKNKTVS---KLVIQANVANSAMYK-----CVVSNKVQODERLIYFYVTTIPQFT 559
DB 459 --VKIQNSSVSPFGKLVQSSIDYSAPKNGTVCECRAYNDVGKSSAFFNF----- 506

QY 560 IESKPSLELLEGGPVLLSCQADSYKYEHLRWYRLNLSTLHDAHGNPLLLDCKNVHLEFATP 619
DB 507 ----- 506

QY 620 LAASLEEVAPGARHATLSLIPRAPEHGHYVCEVQDRRSHDKHCHKYLSVQALEAPR 679
DB 507 --AFKEQI-----HPTLF----- 518

QY 680 LTQNLTDLLVNVSDSLEMQCLVAGAHAPSIVWYKDERLLEKSGVDLADSNQKLSQVRV 739
DB 519 -----TPLLIGSVVAGMCMVIV-----MVLTYK----- 542

QY 740 EEDAGRYLCSVCAKGCNVSSASVAVGSESDKGSMEIVILVGTGVIATVFFWLLLLIFCN 799
DB 543 -----YL-----QKPMYEVQ-----WKVVEI--- 559

QY 800 MRRPAHADITGYLSIIMDPGEVPLEEQEYLSYDASQNEFPRELRHLGVLGYGAFKV 859
DB 560 -----NGNNYVIDPTOLP-----YD-HKWEFPRNRLSFGKTLGAGAFKV 599

QY 860 VEASAFGIHGSSCDTVAVKMLKEGATASEHRALMSELKILIHIGNHLNVNLLGACTKP 919
DB 600 VEATAYGLIKSDAANTVAVKMLKPSAHLTEREALMSELKVLSTYLGNNHNVNLLGACT-V 658

QY 920 QGPLMVIPECKYGNLSNFLRAKRDAPSCAEKSPQGRGRFRAWELAR-----L 969

DB 659 GGPTLVITEYCCYCGDLLNFRKRKDSFCSKQEDHAEALYKNLLHKSCESSNDSTNEYM 718
QY 970 DRRRFGSSDRVLFARFSKTEGGARRA-----SPD-OEADLWLSPLTMEDLVY 1017
DB 719 D-MKFCVSYV-----PTKDKRRAARIGSYIERDVAPSIMEDDEL---ALDLEDLLSF 768

QY 1018 SFQVARGMEFLASRKCIIHRDLAARNILLSQDVVKICDFGLARDIYKDPDYVRKGSARLP 1077
DB 769 SYQVAKGMAFLASKNCIIHRDLAARNILLTHGRYKICDFGLARDIKNDSNYVVGKGNVRLP 828

QY 1078 LKMAPESEIFDKVYTTQSDVMSFGVLLWEIFSLGASPPYGVQINEEFQORLDRGTRMRAP 1137
DB 829 VKMAPESEIFNCVYTFESDVMSYGFILWELFSLGSSPYGHPVDSKFKYMIKEGFRMLSP 888

QY 1138 ELATPAIRIMLNCWSDGPKAPAFSELVEIL 1169
DB 889 EHAPAEYDINKTCWDADPLKRPTFKQIVOLI 920

Search completed: January 23, 2001, 05:05:52
Job time: 210 sec

XX WPI: 1995-366160/47.
DR P-PSDB: R85937.
XX
XX Agonist antibodies which activate specific protein tyrosine
PT kinase(s) - also activate chimeric proteins of kinase extracellular
PT domain and Ig constant domain, useful for studying, and therapeutic
PT modulation of, cell growth and differentiation
XX
XX Disclosure: Fig 15A-F: 125pp; English.
XX
XX DNA probes based on protein tyrosine-kinase (PTK) sequences were used
CC to screen cDNA libraries to identify novel PTK genes. The SAL-S1
CC gene (see also T03101) was isolated from several megakaryocyte cell
CC libraries. The gene can be used to produce recombinant SAL-S1 or its
CC fragments, to detect related genes, and to design drugs, peptides
CC or antisense nucleotides that modulate PTK activity.
XX
XX Sequence 4425 BP; 939 A; 1348 C; 1361 G; 777 T; 0 other;
SQ

Query Match 95.2%; Score 3912.4; DB 16; Length 4425;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3913; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ccacgcgacgcccgcgagatgacgagggcgccgctgtgctgcgactgtgctctg 60
DB 12 ccacgcgacgcccgcgagatgacgagggcgccgctgtgctgcgactgtgctctg 71
QY 61 cctgggactcctggaagcctgtgagtgagtgactcctatgaccccccgcacattaaacat 120
DB 72 cctgggactcctggaagcctgtgagtgagtgactcctatgaccccccgcacattgaaacat 131
QY 121 cagcagagatgcacagtcacacgacccggtgacgctgtcctatcctgcaggagaca 180
DB 132 cagcagagatgcacagtcacacgacccggtgacgctgtcctatcctgcaggagaca 191
QY 181 gacccccctcagtgaggcttggcagagagctcaggagggcgccagccacccggagacaaga 240
DB 192 gacccccctcagtgaggcttggcagagagctcaggagggcgccagccacccggagacaaga 251
QY 241 cagcagagacaggggtgtgagagactgcagagcagcagagccgctacttgcaa 300
DB 252 cagcagagacaggggtgtgagagactgcagagcagcagagccgctacttgcaa 311
QY 301 ggtgtgtcgtgcacaggtacatgccaacgacacagcagcagctgctgtactacaa 360
DB 312 ggtgtgtcgtgcacaggtacatgccaacgacacagcagcagctgctgtactacaa 371
QY 361 gtacatcaaggcacgacatcagagccacacgcccgcagctcctacgtgtcgtgagaga 420
DB 372 gtacatcaaggcacgacatcagagccacacgcccgcagctcctacgtgtcgtgagaga 431
QY 421 ctttgagcagcattatcaacaagcctgacacgcttcttggtcaacaggaagacgcat 480
DB 432 ctttgagcagcattatcaacaagcctgacacgcttcttggtcaacaggaagacgcat 491
QY 481 gtgggtgcccctgtgtgttcctatccccgcctcaatgtacgctgcgctcgcaaaagctc 540
DB 492 gtgggtgcccctgtgtgttcctatccccgcctcaatgtacgctgcgctcgcaaaagctc 551
QY 541 ggtgtgtgtggcagagggcagggaggtgtgtgtgagatgacccggcggggacgtcgtgtc 600
DB 552 ggtgtgtgtggcagagggcagggaggtgtgtgtgagatgacccggcggggacgtcgtgtc 611
QY 601 cagcgcactgtgcacagatccctgtacctcagtgagagaccacccctgggagaccagga 660
DB 612 cagcgcactgtgcacagatccctgtacctcagtgagagaccacccctgggagaccagga 671
QY 661 cttcctttccaaacccctcctcctgtgtgcacatcacaggcaacgagctctatgacatccagct 720
DB 672 cttcctttccaaacccctcctcctgtgtgcacatcacaggcaacgagctctatgacatccagct 731

QY 721 gttgccaggaagtcgctgagagctgctgtgaggggagagagctggtcctgaactgacacct 780
DB 732 gttgccaggaagtcgctgagagctgctgtgaggggagagagctggtcctgaactgacacct 791
QY 781 gtgggtgagctttaactcaggttcaccttgcagtgagactcagcaggaagcagcagaga 840
DB 792 gtgggtgagctttaactcaggttcaccttgcagtgagactcagcaggaagcagcagaga 851
QY 841 gcggggttaagtgggtgcccgcgagcgcgtcccagcagacccacacagaaactctccagcat 900
DB 852 gcggggttaagtgggtgcccgcgagcgcgtcccagcagacccacacagaaactctccagcat 911
QY 901 cctgaccatccacaacgctcagcagacgacctgggctcgtatgtgcaagcccaacaa 960
DB 912 cctgaccatccacaacgctcagcagacgacctgggctcgtatgtgcaagggccaacaa 971
QY 961 cggcatccaggtattcgggagagcaccgaggtcattgtgcatgaaatcccttcatcag 1020
DB 972 cggcatccaggtattcgggagagcaccgaggtcattgtgcatgaaatcccttcatcag 1031
QY 1021 cgtcagtggtctcaaaagacccatccttggggccacgcccagggagacgagctggtgaagct 1080
DB 1032 cgtcagtggtctcaaaagacccatccttggggccacgcccagggagacgagctggtgaagct 1091
QY 1081 gccctaaagctggcagcgtaccctcccccgcaggttccagtggtacaagatggaagggc 1140
DB 1092 gccctaaagctggcagcgtaccctcccccgcaggttccagtggtacaagatggaagggc 1151
QY 1141 actgtccggggcgccacagtcacatgcccctggtgctcaagagagtgacagagggccagcac 1200
DB 1152 actgtccggggcgccacagtcacatgcccctggtgctcaagagagtgacagagggccagcac 1211
QY 1201 aggcacttaacccctcctcgtgaaactcgtgctgctgagcagggcgcaacatcagct 1260
DB 1212 aggcacttaacccctcctcgtgaaactcgtgctgctgagcagggcgcaacatcagct 1271
QY 1261 aggcctggtgtaagtgtgccccccagatacatgagaagagggcctcctccccagcat 1320
DB 1272 aggcctggtgtaagtgtgccccccagatacatgagaagagggcctcctccccagcat 1331
QY 1321 ctactcgcgtcacagccgcccagcctcactgcacgctcagcgggtgccccctgcctct 1380
DB 1332 ctactcgcgtcacagccgcccagcctcactgcacgctcagcgggtgccccctgcctct 1391
QY 1381 cagcatccagtgccactgcccgcctggacacccctgcaagatgtttgccagcgtagctct 1440
DB 1392 cagcatccagtgccactgcccgcctggacacccctgcaagatgtttgccagcgtagctct 1451
QY 1441 ccggcgccggcagcagcagacccatcagagcctggacacccctggacccaggtttgtggagggaaa 1500
DB 1452 ccggcgccggcagcagcagacccatcagagcctggacacccctggacccaggtttgtggagggaaa 1511
QY 1501 gcaagatgcccgtgaacccccatcagagcctggacacccctggacccaggtttgtggagggaaa 1560
DB 1512 gcaagatgcccgtgaacccccatcagagcctggacacccctggacccaggtttgtggagggaaa 1571
QY 1561 gaataagactgtgagcaagctggtgatccagaatgccagatgctgctgcatcacaagt 1620
DB 1572 gaataagactgtgagcaagctggtgatccagaatgccagatgctgctgcatcacaagt 1631
QY 1621 gttggtctcaacaaagtgggccagagatgagcggctcactacttctatgtgaccacat 1680
DB 1632 gttggtctcaacaaagtgggccagagatgagcggctcactacttctatgtgaccacat 1691
QY 1681 ccccgacgcttcaacatccgaatccaaagcctccgagagactactagagggccagcgggt 1740
DB 1692 ccccgacgcttcaacatccgaatccaaagcctccgagagactactagagggccagcgggt 1751
QY 1741 gctcctgagctgccaaagccgacagctacaagtacgagcactctgctggttacgcgctcaa 1800
DB 1752 gctcctgagctgccaaagccgacagctacaagtacgagcactctgctggttacgcgctcaa 1811
QY 1801 cctgtccacgctgcacgactgcgcacggggaacccgcttctgctcagactgcaagaacgtgca 1860

1812 cctgtccacgctgcaagatgacagcggaaacccgcttctgtctgactgcaagaacgctgca 1871
1861 tctgttccgcaacccctctgtccgcaacccgctgagagagtggtgacactggtggcgccgcaagc 1920
1872 tctgttccgcaacccctctgtccgcaacccgctgagagagtggtgacactggtggcgccgcaagc 1931
1921 cagctcagcctgagatcccccgcgtgcgcacagcagcagcagcagcagcagcagcagcagcagc 1980
1932 cagctcagcctgagatcccccgcgtgcgcacagcagcagcagcagcagcagcagcagcagcagc 1991
1981 agtgcgaagcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040
1992 agtgcgaagcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2051
2041 cctggaagccctcggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
2052 cctggaagccctcggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2111
2101 gctggagatgagctgctggtggcgagcagcagcagcagcagcagcagcagcagcagcagcagc 2160
2112 gctggagatgagctgctggtggcgagcagcagcagcagcagcagcagcagcagcagcagcagc 2171
2161 cgagagctgctggagagagagctgtgagctgagctgagcagcagcagcagcagcagcagcagc 2220
2172 cgagagctgctggagagagagctgtgagctgagcagcagcagcagcagcagcagcagcagcag 2231
2221 catccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2280
2232 catccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2291
2281 gggctgcgtcaactctccgcaagcgtggtggcgagcagcagcagcagcagcagcagcagcagc 2340
2292 gggctgcgtcaactctccgcaagcgtggtggcgagcagcagcagcagcagcagcagcagcagc 2351
2341 gggagatgagctgctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400
2352 gggagatgagctgctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2411
2401 cctcatctctgtaacatgagagagcgcgcgcacgagcagcagcagcagcagcagcagcagcagc 2460
2412 cctcatctctgtaacatgagagagcgcgcgcacgagcagcagcagcagcagcagcagcagcagc 2471
2461 catcatatggaccccgaggagctgctctgagagagcagcagcagcagcagcagcagcagcagcagc 2520
2472 catcatatggaccccgaggagctgctctgagagagcagcagcagcagcagcagcagcagcagcagc 2531
2521 tgcagcagctggaattcccccagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2580
2532 tgcagcagctggaattcccccagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2591
2581 cgcttcgggaaagtggtggaagcctcctgcttccgcatccacaaagggcagcagcagcagcagcagc 2640
2592 cgcttcgggaaagtggtggaagcctcctgcttccgcatccacaaagggcagcagcagcagcagcagc 2651
2641 caccgtgagcagtgaaatgctgaaagagcgcgcacagcagcagcagcagcagcagcagcagcagcagc 2700
2652 caccgtgagcagtgaaatgctgaaagagcgcgcacagcagcagcagcagcagcagcagcagcagcagc 2711
2701 gtcgagctcaagatcctcatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2760
2712 gtcgagctcaagatcctcatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2771
2761 ggcgtgaccaaagccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2820
2772 ggcgtgaccaaagccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2831
2821 cctctcaactctcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2880
2832 cctctcaactctcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2891
2881 cgagc 2940

2892 cgagc 2951
2941 gggagc 3000
2952 gggagc 3011
3001 ggccttccgacacaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3060
3012 ggccttccgacacaaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3071
3061 gctctgctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3120
3072 gctctgctacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3131
3121 ccacagagacacctgctctcggaacatctctgctcggaagcagcagcagcagcagcagcagcagcagc 3180
3132 ccacagagacacctgctctcggaacatctctgctcggaagcagcagcagcagcagcagcagcagcagc 3191
3181 tgaacttgacctgctcgcgagacatctacaaagacccctgactacgtcccgcaagggcagtgac 3240
3192 tgaacttgacctgctcgcgagacatctacaaagacccctgactacgtcccgcaagggcagtgac 3251
3241 ccgagc 3300
3252 ccgagc 3311
3301 gagtgagc 3360
3312 gagtgagc 3371
3361 gtacactggggtgagagatcaatgagagagcttctgagagagcagcagcagcagcagcagcagcagcagc 3420
3372 gtacactggggtgagagatcaatgagagagcttctgagagagcagcagcagcagcagcagcagcagcagc 3431
3421 gaggcccgagc 3480
3432 gaggcccgagc 3491
3481 agaccccaagc 3540
3492 agaccccaagc 3551
3541 gggcaggggctgcaagaggaagagagctgcatggcccgagcagcagcagcagcagcagcagcagcagc 3600
3552 gggcaggggctgcaagaggaagagagctgcatggcccgagcagcagcagcagcagcagcagcagcagc 3611
3601 agaaagggcagc 3660
3612 agaaagggcagc 3671
3661 tagagc 3720
3672 tagagc 3731
3721 gtccttcccgagc 3780
3732 gtccttcccgagc 3791
3781 atttgaggatcccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3840
3792 atttgaggatcccccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3851
3841 cagtgagatggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3900
3852 cagtgagatggctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3911
3901 aagcggctcagcagc 3914
3912 aagcggctcagcagc 3925

ID T03104 standard; DNA; 9108 BP.
XX
AC T03104:
XX
XX
DT 14-FEB-1996 (first entry)
XX
DE Plasmid pRK5.tk1-1.1 encoding Sal S-1.
XX
XX Protein tyrosine-kinase; pTK; SAL-S1; agonist; cell growth;
KW differentiation; pRK5.tk1-1.1; ss.
XX
XX Chimeric synthetic;
OS Chimeric Homo sapiens.
XX
XX W09527061-AL.
XX
XX 12-OCT-1995.
XX
PD 04-APR-1995; 95WO-US04228.
XX
PF 04-APR-1994; 94US-0222616.
XX
PR (GETH) GENENTECH INC.
XX
XX Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI Wood WI;
PI
XX MPI; 1995-366160/47.
DR
XX Agonist antibodies which activate specific protein tyrosine
XX kinase(s) - also activate chimeric proteins of kinase extracellular
PT domain and Ig constant domain, useful for studying, and therapeutic
PT modulation of, cell growth and differentiation
XX
XX Example 7; Page 102-111; 125pp; English.
XX
XX PCR primers given in T03102-03 were used to amplify plasmid
CC pRK5.tk1-1.1 (T03104) incorporating SAL-S1 (see T03090) DNA.
CC The product was used to construct a fusion of the SAL-S1
CC extracellular domain and human IgG1 FC domain, which was expressed
CC in 293 and COS7 cells and used to raise antibodies having
CC protein tyrosine-kinase agonist activity.
XX
XX Sequence 9108 BP; 2086 A; 2554 C; 2491 G; 1977 T; 0 other;
SQ

Query Match 95.2%; Score 3912.4; DB 16; Length 9108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3913; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ccacgcgcagcggccgagatgacggggcgccgcgcgtgctgcgcgtgctgctg 60
Db 944 ccacgcgcagcggccgagatgacggggcgccgcgcgtgctgcgcgtgctg 1003
Qy 61 cctgggactcctgacggcctggtgagtgctgactcctcctgaccccccgccttgaaat 120
Db 1004 cctgggactcctgacggcctggtgagtgctgactcctcctgaccccccgccttgaaat 1063
Qy 121 caccgagagatcacagctacgacacccggtgacgctgctcctcctgacggggaca 180
Db 1064 caccgagagatcacagctacgacacccggtgacgctgctcctcctgacggggaca 1123
Qy 181 gaccccccctgagtgggcttggcagagatcagggcgccagggccacccggagacaagga 240
Db 1124 gaccccccctgagtgggcttggcagagatcagggcgccagggccacccggagacaagga 1183
Qy 241 caccgagacacggggtggtgagagatcagggcgacagacgacggccctactgcaa 300
Db 1184 caccgagacacggggtggtgagagatcagggcgacagacgacggccctactgcaa 1243
Qy 301 ggtattgctgctgacgaggtacatgccacgacacacagggcagctgctgctactacaa 360
Db 1244 ggtattgctgctgacgaggtacatgccacgacacacagggcagctgctgctactacaa 1303

Qy 361 gtacatcaggcagcatcagggcaccacggcccgccagctcctacgtgttcgtgagaga 420
Db 1304 gtacatcaggcagcatcagggcaccacggcccgccagctcctacgtgttcgtgagaga 1363
Qy 421 ctttgaagccattcctcaacaaagctgacagcgtcttggtaaacaggaagcagcat 480
Db 1364 ctttgaagccattcctcaacaaagctgacagcgtcttggtaaacaggaagcagcat 1423
Qy 481 gtgggtgcccgtgctgctgctcctcaatgctacgctgcgctgcgcaagctc 540
Db 1424 gtgggtgcccgtgctgctgctcctcaatgctacgctgcgctgcgcaagctc 1483
Qy 541 ggtgctgaggccagcgggaggggtggtggtggtgacccggcggggcatgctcgtgctc 600
Db 1484 ggtgctgaggccagcgggaggggtggtggtggtgacccggcggggcatgctcgtgctc 1543
Qy 601 caccgcaatgctgacagatgcccgtgacctgacgtgacgacacccctggggagaccagga 660
Db 1544 caccgcaatgctgacagatgcccgtgacctgacgtgacgacacccctggggagaccagga 1603
Qy 661 ctctcttccaaaccccttccgtgacatcacagggcaacgagctctatgacatccagct 720
Db 1604 ctctcttccaaaccccttccgtgacatcacagggcaacgagctctatgacatccagct 1663
Qy 721 gttgcccagggaagtgcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 780
Db 1664 gttgcccagggaagtgcgtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 1723
Qy 781 gtgggtgagtttaactcagatgctcacttgcactgggactaccacaggaagcagcagga 840
Db 1724 gtgggtgagtttaactcagatgctcacttgcactgggactaccacaggaagcagcagga 1783
Qy 841 gcgggttaagtgggtgcccgcgacgctcccagcagacccacacagaaactctccagcat 900
Db 1784 gcgggttaagtgggtgcccgcgacgctcccagcagacccacacagaaactctccagcat 1843
Qy 901 cctgaccatccaaacgtcagccagcagcagcagcagcagcagcagcagcagcagcagcagc 960
Db 1844 cctgaccatccaaacgtcagccagcagcagcagcagcagcagcagcagcagcagcagcagc 1903
Qy 961 cggcattccagcagatttcgggagcagcagcagcagcagcagcagcagcagcagcagcagc 1020
Db 1904 cggcattccagcagatttcgggagcagcagcagcagcagcagcagcagcagcagcagcagc 1963
Qy 1021 cgtcagatggctcaaaagaccatcctggaggccacgcgagagacgagcagcagcagcagcagc 1080
Db 1964 cgtcagatggctcaaaagaccatcctggaggccacgcgagagacgagcagcagcagcagcagc 2023
Qy 1081 gcccgtgaagctggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
Db 2024 gcccgtgaagctggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2083
Qy 1141 actgtccgggcccacagctccacatgcccctggtgctcaaggaggtgacagagccagcagc 1200
Db 2084 actgtccgggcccacagctccacatgcccctggtgctcaaggaggtgacagagccagcagc 2143
Qy 1201 aggcacctacacccctgcgcctgtgaaactccgctgctgctgctgctgctgctgctgctgctgct 1260
Db 2144 aggcacctacacccctgcgcctgtgaaactccgctgctgctgctgctgctgctgctgctgctgct 2203
Qy 1261 ggaagctgggtggaatgtgccccccagatatacaggaagggagggcctctccccagcat 1320
Db 2204 ggaagctgggtggaatgtgccccccagatatacaggaagggagggcctctccccagcat 2263
Qy 1321 ctactgcgtcagcgcgcagcgcgcagcgcgcagcgcgcagcgcgcagcgcgcagcgcgcctct 1380
Db 2264 ctactgcgtcagcgcgcagcgcgcagcgcgcagcgcgcagcgcgcagcgcgcagcgcgcctct 2323
Qy 1381 cagcatccagtgagcagc 1440
Db 2324 cagcatccagtgagcagc 2383

QY 1441 ccggcgagcagcagcaaacctcatgcccacagtccgtgactggaggcggtgacccac 1500
|||||
Db 2384 ccggcgcgccgagcagcaaacctcatgcccacagtccgtgactggaggcggtgacccac 2443
QY 1501 gcaggatgccgtgaaccccatcgagagcctggacacctggaccaggtttgtggaggaaaa 1560
|||||
Db 2444 gcaggatgccgtgaaccccatcgagagcctggacacctggaccaggtttgtggaggaaaa 2503
QY 1561 gaaataagactgagcaagctggatgccagaatgccaacgtgtctccatgtacaagtg 1620
|||||
Db 2504 gaaataagactgagcaagctggatgccagaatgccaacgtgtctccatgtacaagtg 2563
QY 1621 tgtgtctcccaacagggtggggccaggatgagcgctcatctacttctatgtgacccact 1680
|||||
Db 2564 tgtgtctcccaacagggtggggccaggatgagcgctcatctacttctatgtgacccact 2623
QY 1681 ccccgagcggttcacactcgaaatccaagcraatccgagagctactagagccacgcggt 1740
|||||
Db 2624 ccccgagcggttcacactcgaaatccaagcraatccgagagctactagagggccagcggt 2683
QY 1741 gctcctgagctgccaagccgacagctacaagtacgagcatctgctgtgttaccgcctcaa 1800
|||||
Db 2684 gctcctgagctgccaagccgacagctacaagtacgagcatctgctgtgttaccgcctcaa 2743
QY 1801 cctgtccacgctgcacgatgcgcagcggaacccgcttctgtctgactgcaagaacgtgca 1860
|||||
Db 2744 cctgtccacgctgcacgatgcgcagcggaacccgcttctgtctgactgcaayaacgtgca 2803
QY 1861 tctgttcgcacccctctggccgcagccttgagagaggtggcacctggggcgccacgc 1920
|||||
Db 2804 tctgttcgcacccctctggccgcagccttgagagaggtggcacctggggcgccacgc 2863
QY 1921 cacgctcaagcctgagatccccccgctgcgcgcccgagcacagggccactatgtgtgcga 1980
|||||
Db 2864 cacgctcaagcctgagatccccccgctgcgcgcccgagcacagggccactatgtgtgcga 2923
QY 1981 agtgcagaacccgagcagcctgacaagcactgccacaagaagtacctgtcggtgcagcc 2040
|||||
Db 2924 agtgcagaacccgagcagcctgacaagcactgccacaagaagtacctgtcggtgcagcc 2983
QY 2041 cctggaagccctcggtctacgcagagaacttgaccgacctctctgttgaaagtgcagcact 2100
|||||
Db 2984 cctggaagccctcggtctacgcagagaacttgaccgacctctctgttgaaagtgcagcact 3043
QY 2101 gctggagatgcaagtctgtgagccgagcgacgcgcgcacacatcgltgtgtacaaga 2160
|||||
Db 3044 gctggagatgcaagtctgtgagccgagcgacgcgcgcacacatcgltgtgtacaaga 3103
QY 2161 cgagagcctgctggagaaaaactctggagctgacttggcgactcccaaccagaagctgag 2220
|||||
Db 3104 cgagagcctgctggagaaaaactctggagctgacttggcgactcccaaccagaagctgag 3163
QY 2221 catccagcgctgcgcagagagatgagggacgctatctgtgcagcgtgtgcaacgcaa 2280
|||||
Db 3164 catccagcgctgcgcagagagatgagggacgctatctgtgcagcgtgtgcaacgcaa 3223
QY 2281 gggctcgctcaactctccgcagagctggcgtggagagctccgaggaataaggagcagat 2340
|||||
Db 3224 gggctcgctcaactctccgcagagctggcgtggagagctccgaggaataaggagcagat 3283
QY 2341 gggagatgctgactctgtcgggtacccgctcatcgctgtcttcttctgtgggtctctct 2400
|||||
Db 3284 gggagatgctgactctgtcgggtacccgctcatcgctgtcttcttctgtgggtctctct 3343
QY 2401 cctcatctctgtaacatgagagggccgcccacgcagacatcaagaagggttacctgtc 2460
|||||
Db 3344 cctcatctctgtaacatgagagggccgcccacgcagacatcaagaagggttacctgtc 3403
QY 2461 catcatatggaccccgaggaggtgctctgtgaggagcaatgcgaatacctgtctacga 2520
|||||
Db 3404 catcatatggaccccgaggaggtgctctgtgaggagcaatgcgaatacctgtctacga 3463
QY 2521 tggcagcagtggaattcccccgagagcggtctgcacctgggggagagtgctcggttacgg 2580
|||||

Db 3464 tggcagcagtggaattcccccgagagcggtgcacctgggggagagtgctcggttacgg 3523
QY 2581 cgctctcggaaggttggtggaagcctccgcttccgcatccacaaggcgacgagctgtga 2640
|||||
Db 3524 cgctctcggaaggttggtggaagcctccgcttccgcatccacaaggcgacgagctgtga 3583
QY 2641 caccgtggcgctgaaatgctgaaagaggcgccacggcgacgagcagccgcgctgat 2700
|||||
Db 3584 caccgtggcgctgaaatgctgaaagaggcgccacggcgacgagcagccgcgctgat 3643
QY 2701 gtcggagctcaagatctcatctacatcgccaacacctcaacgtgtgtaacctctccgg 2760
|||||
Db 3644 gtcggagctcaagatctcatctacatcgccaacacctcaacgtgtgtaacctctccgg 3703
QY 2761 ggcgtgcaccaagccgagggcgccctcatggtgatgctggagttctgcaagtacggcaa 2820
|||||
Db 3704 ggcgtgcaccaagccgagggcgccctcatggtgatgctggagttctgcaagtacggcaa 3763
QY 2821 cctctcaaatctcctgcgcgaagcgagcgcttcagccctcgcgcgagagagctcc 2880
|||||
Db 3764 cctctcaaatctcctgcgcgaagcgagcgcttcagccctcgcgcgagagagctcc 3823
QY 2881 cgagcagcgagcagctccgcgcctatgggtgagctcgccaggtgagcgagcgccg 2940
|||||
Db 3824 cgagcagcgagcagctccgcgcctatgggtgagctcgccaggtgagcgagcgccg 3883
QY 2941 ggggagcagcagcagggctcctcttcgcgcggttctcgaaagaccgagggcgagcgagcg 3000
|||||
Db 3884 ggggagcagcagcaggggtcctcttcgcgcggttctcgaaagaccgagggcgagcgagcg 3943
QY 3001 ggccttccagaccaaagaagctgagggacctgtgctgagcccgctgacctggaagatct 3060
|||||
Db 3944 ggccttccagaccaaagaagctgagggacctgtgctgagcccgctgacctggaagatct 4003
QY 3061 tgtgtctacagcttcagggtggcgagaggtgaggttctcggttcccgaaagtgat 3120
|||||
Db 4004 tgtgtctacagcttcagggtggcgagaggtgaggttctcggttcccgaaagtgat 4063
QY 3121 ccagagagcctgctgctcgaaacattctgctgcgaaagcgacgtgggtgaaagatctg 3180
|||||
Db 4064 ccagagagcctgctgctcgaaacattctgctgcgaaagcgacgtgggtgaaagatctg 4123
QY 3181 tgactttggccttgcggggacatctacaaagacctgactacgtccgcaaggcgagtg 3240
|||||
Db 4124 tgactttggccttgcggggacatctacaaagacctgactacgtccgcaaggcgagtg 4183
QY 3241 ccggtgcacctgagtgagtgagccctgaaagcatcttcgacaaggtgtacaccagca 3300
|||||
Db 4184 ccggtgcacctgagtgagtgagccctgaaagcatcttcgacaaggtgtacaccagca 4243
QY 3301 gagtgcagtggtgctcttgggggtgcttctctggagagatcttctctgggggctcccc 3360
|||||
Db 4244 gagtgcagtggtgctcttgggggtgcttctctggagagatcttctctgggggctcccc 4303
QY 3361 gtacctgggggtgcagatacaatgagaggttctgcaggggctgagagcgcaagaagat 3420
|||||
Db 4304 gtacctgggggtgcagatacaatgagaggttctgcaggggctgagagcgcaagaagat 4363
QY 3421 gaggggcccgagctggcgacctcccgccatagccgcgcatctgctgaaactgctgggtccgg 3480
|||||
Db 4364 gaggggcccgagctggcgacctcccgccatagccgcgcatctgctgaaactgctgggtccgg 4423
QY 3481 agaccaccaaggcgagacctgcattctcgagctgtgtgagatctctgggggacctctcca 3540
|||||
Db 4424 agaccaccaaggcgagacctgcattctcgagctgtgtgagatctctgggggacctctcca 4483
QY 3541 gggcaggggctgcagaagagagaggtctgcatggcccgcgacgctctcagagctc 3600
|||||
Db 4484 gggcaggggctgcagaagagagagaggtctgcatggcccgcgacgctctcagagctc 4543
QY 3601 agaaaggggagctctctcgaggtgtccaccatggccctacacatcgcccgaggtgacgc 3660
|||||


```
QY 3146 attctcgtcgcgaagagcagctggtgaagatctgtgaacttggccttcccggaacatc 3205
Db 2647 ATTCTGCTGCGGAAAGCAGCTGGTGAAGATCTGTACATTTGGCCTTCCCGGGACATC 2588
QY 3206 tacaagaccctgaactacgtccgaagagcagtgccgagctgccctgaagtgaatgcc 3265
Db 2587 TACAAGACCCGACTACGTCCGAAGGAGTGCCTCCGCTGAGGTGGATGGCC 2528
QY 3266 cctgaagcatcttcgacaaggtgtacaccacgcagagtgacgtgtggtcctttggggg 3325
Db 2527 CCTGAAAGCATCTTCACAAAGGTGTACACCACGACAGTGAGTGTGCTTGGGGTG 2468
QY 3326 ctctctggaagatcttctctctggtgggctcccccacccctgagatcaatgag 3385
Db 2467 CTCTCTGGGAGATCTCTCTCTGGGGGCTCCCTACCTCTGGGGTGAGATCAATGAG 2408
QY 3386 gagttctgcagcgcgtgagagcgcacaaagatgagggcccgagctggccatccc 3445
Db 2407 GAGTTCTGCAGCGGTGAGAGAGCGCACAAAGATGAGGGCCCGGAGGTGGCCACTCC 2348
QY 3446 gccatacgcgcacatctgctgaactgctggtccggagaccccaaggcgagacctgcattc 3505
Db 2347 GCCATACGCCGCATCATGCTGAATCTGTGTCGGGAGACCCCAAGCGAGACTGCTCATTC 2288
QY 3506 tcgagactggtgagatcctggggagcctgctccagggcagggcctgcaagaggaag 3565
Db 2287 TCGGAGCTGGTGAGATCTCTGGGGGACCTGCTCCAGGGCAGGGCCCTGCAAGAGGAAG 2228
QY 3566 gaggctgcagtgcccgagcagctctcagagctcagaagagggcagctctcgcaggtg 3625
Db 2227 GAGSTCTGCATGSCCGCGCAGCTCTCAGAGCTCAGAGGTCAGAGAGGGCAGCTTCTCGCAGGTG 2168
QY 3626 tcacacatggcctcacatctgcccaggtgagctgagagacagcccgcaagcctgcag 3685
Db 2167 TCCACCATGGCCCTAGACATTCGCCAGGTGAGCTCAGGACAGCCCGCAAGCCCTGCAG 2108
QY 3686 cgcacagcctgcccaggtattacaactgggtgctcttcccggtgcttgcctgagcaga 3745
Db 2107 CGCCACAGCTGCCGCCAGGTATTACAACTGGGTGCTCTTTCCCGGTGCTTGGCCAGA 2048
QY 3746 ggggctgagaccctggttctcctcaggtgagagacatttgaggaattcccccagcccca 3805
Db 2047 GGGGCTGAGACCCGTGTTCTCTCAGGATGAAGACATTTGAGGAATTCCTCCATGACCCCA 1988
QY 3806 acgacctataaagctctgtgacaaccagacagcagtgaggatgctgctgctcggag 3865
Db 1987 ACGACCTACAAAGGCTCTGTGGACAACACAGACAGTGGGATGGTGTGCTGGCCTCGGAG 1928
QY 3866 gagtttgacagatagagcagagcatagacaagaagcgggttcagct 3914
Db 1927 GAGTGTGAGCAGATAGAGCAGGTATAGACAAGAAGCGGCTTCAGGT 1879
```

RESULT 5

```
T03101/c
ID T03101 standard; DNA; 6827 BP.
XX
AC T03101;
XX
DT 14-FEB-1996 (first entry)
XX
DE Protein tyrosine-kinase SAL-S1 gene.
XX
KW Protein tyrosine-kinase; pTK; SAL-S1; agonist; cell growth;
KW differentiation; ss.
XX
OS Homo sapiens.
XX
FH .Key Location/Qualifiers
FT misc_difference 3026
FT /*tag= a
FT /note= "base n at position 3026 is not identified
FT in the specification"
```

```
XX W09527061-AL.
XX 12-OCT-1995.
XX
PF 04-APR-1995; 95WO-US04228.
XX
PR 04-APR-1994; 94US-0222616.
XX
PA (GETH ) GENENTECH INC.
XX
PI Bennett BD, Goeddel D, Lee JM, Matthews W, Tsai SP;
PI Wood WI;
XX
XX WPI: 1995-366160/47.
XX P-PSDB; R85938.
XX
XX Agonist antibodies which activate specific protein tyrosine
XX kinase(s) - also activate chimeric proteins of kinase extracellular
XX domain and Ig constant domain, useful for studying, and therapeutic
XX modulation of, cell growth and differentiation
XX
XX Disclosure: Page 39-46; 125pp; English.
XX
XX DNA probes based on protein tyrosine-kinase (pTK) sequences were used
XX to screen cDNA libraries to identify novel pTK genes. The SAL-S1
XX gene (see also 103090) was isolated from several megakaryocytic cell
XX libraries. The gene can be used to produce recombinant SAL-S1 or its
XX fragments, to detect related genes, and to design drugs, peptides
XX or antisense nucleotides that modulate pTK activity.
XX
XX Sequence 6827 BP; 1558 A; 1870 C; 1738 G; 1660 T; 1 other;
XX
XX
XX Query Match 25.3%; Score 1040.2; DB 16; Length 6827;
XX Best Local Similarity 98.3%; Pred. No. 3.4e-200;
XX Matches 1051; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 2846 cgggagccttcagccctgcgaggaagtctcccgagcagcgagccttcgcgcgc 2905
Db 2947 CGAGCCAGCTTCGTGCATCCGCGAGGAGAGTCTCCCGAGCAGCGGAGCTTCCGCGCC 2888
QY 2906 atggtggaagctccagcagctgagtcgagcgagcgagcagcagcagctctcttc 2965
Db 2887 ATGTTGAGCTCCCGAGGCTGATCGAGCGCGCGGAGGAGCGACAGGCTCTCTTC 2828
QY 2966 gcgcggttctcgaagcagcgagcgagcgagcgagcgagcgagcgagcgagcgagc 3025
Db 2827 GCGCGGTTCTCGAAGACCGAGGCGGAGCGAGGCGGCTTCTCCAGACCAAGACTGAG 2768
QY 3026 gacctgagctgagcccgctgacctgaagatcttctgtctacagcttccagtgccc 3085
Db 2767 GACCTGTGGCTGAGCCCGCTGACCATGGAAGATCTGTGTACAGCTTCCAGGTGCC 2708
QY 3086 agagggatgagttctctggttcccgaaagtgcattccacagagcctgagcctcggaac 3145
Db 2707 ACAGGGATGGAGTTCTTGGCTTCCGAAAGTGCATCCACAGAGAGACTGGCTCGGAAC 2648
QY 3146 attctgctggaagcagcagtggtgaagatctgtgactttgacctgcccgggacatc 3205
Db 2647 ATTCTGCTGCGAAGGAGCAGCTGGTGAAGATCTGTGACTTTGGCCTTCCCGGGACATC 2588
QY 3206 tacaagaccctgactacgtccgaagagcagtgcccgagctgccctgaagtgaatgcc 3265
Db 2587 TACAAGACCCCGCTACTACTCTCCGAAAGGAGTGCCTCCGCTGCCCTTGAAGTGGATGGCC 2528
QY 3266 cctgaagcatcttcgacaaggtgtacaccacgcagagtgacgtgtggtcctttggggg 3325
Db 2527 CCTGAAAGCATCTTCGACAAGGTGTACACCACGAGAGTGAGCTGTGCTTCCGCTG 2468
QY 3326 ctctctggaagatcttctctctggtgggctcccccacccctgagatcaatgag 3385
Db 2467 CTCTCTGGGAGATCTCTCTCTGGGGGCTCCCTACCTGAGGCTGCGATCAATGAG 2408
```

QY 3386 gagttctgccaggcgtgagagcagcagagatgagggcccgagcgtgccaactccc 3445
DB 2407 GAGTCTTCCAGCGCTGAGAGCGGCAAGAGTAGGGCCCGGAGCTGGCCACTCCC 2348
QY 3446 gccatccgcgcacatcatgaactcgtcgtcccgagagaccccaagcgagacclqcatc 3505
DB 2347 GCCATACGCCGCATCATGCTGAACCTGTGTCGGAGACCCCAAGCCGAGACCTGCATT 2288
QY 3506 tcggagctggtgagatcctggggacacctgctccagggcagggcctgcaagaggaag 3565
DB 2287 TCGGAGCTGCTGGAGATCCTGGGGACCTGTCTCCAGGCGAGGGCTGCAAGAGGAAGAG 2228
QY 3566 gaggtctgatccgcgcagcgtctcagagctcagagagggcagcttctcgcaggtg 3625
DB 2227 GAGGTCTGATGCCCCCGCGAGCTCTCAGAGCTCAGAGAGGGGAGCTTCTCGAGGTG 2168
QY 3626 tcaccatggcctcacatcgtcccgagcgtgacgtgagagcagcccgccagcctgcag 3685
DB 2167 TCCACCATGCCCTACATATGCCAGGCTGACGCTGAGGACAGCCCGCAAGCCTGCAG 2108
QY 3686 cgcacagcctgcccgcagctattacaactgggtcttctcccggtgctgcccaga 3745
DB 2107 CGCCACAGCCTGCCCGCGAGGATTACAACTGGGTGCTCTTCCCGGGTGCCTGGCCAGA 2048
QY 3746 ggggctgagaccctggttctccaggtatgaagacatttgaggaattcccccagccccca 3805
DB 2047 GGGGCTGAGACCCGCTGCTCTCCAGGATGAACACATTTGAGGAATTCGCCCATGACCCA 1988
QY 3806 acgactacaaagctctgtgacaaccagacagacagtgagtgatgctgacctcgag 3865
DB 1987 ACGACCTACAAAGGCTCTGTGGACAACCCAGACAGACAGTGGATGGTGTGCTGGCTCGGAG 1928
QY 3866 gagttgacagatagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3914
DB 1927 GAGTGTGAGCAGATAGAGCAGGATAGACAAAGAAAGCGGCTTCAGGT 1879

RESULT 6
T38735
ID T38735 standard; cDNA; 5404 BP.
XX AC T38735;
XX DT 11-DEC-1996 (first entry)
XX DE Murine foetal liver kinase 1 cDNA.
XX KW Murine: foetal liver kinase 1; flk-1; protein tyrosine kinase;
KW monoclonal; antibody; extracellular domain; receptor assay;
KW haematopoietic stem cell; ligand; stimulation; proliferation;
KW differentiation; treatment; anaemia; bone marrow damage;
KW cancer chemotherapy; radiation; ds.
XX OS Mus musculus.
XX FH Key Location/Qualifiers
FT CDS 208..4311
FT sig_peptide /tag= a
FT FT 208...264
FT FT /*tag= b
FT FT 265...4308
FT FT /*tag= c
XX DN US5548065-A.
XX PD 20-AUG-1996.
XX PF 02-APR-1991; 91US-0679666.
XX PR 19-NOV-1992; 92US-0977451.
XX PR 02-APR-1991; 91US-0679666.
XX PR 28-JUN-1991; 91US-0728913.

PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0906397.
PR 12-NOV-1992; 92US-0975049.
PR 30-APR-1993; 93US-0055269.
PR 31-OCT-1994; 94US-0252517.
XX (UYPV-) UNIV PRINCETON.
XX PA Lemischka IR;
XX PI
XX DR WPI: 1996-392678/39.
XX DR P-PSDB; T38735.
XX PT Anti-foetal liver kinase 2 (flk-2) antibodies - useful in assays,
XX PT for isolating haematopoietic stem cells expressing receptor and for
XX PT obtaining ligands
XX FS Disclosure: Columns 51-62; 50pp; English.
XX PS The present sequence encodes murine foetal liver kinase 1 (flk-1),
XX CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
XX CC raised against the extracellular portion of flk-1 can be used to
XX CC assay for flk receptors on the surface of haematopoietic stem
XX CC cells, and to isolate positive cells. The antibodies can also
XX CC be used as, or to obtain ligands, which stimulate the proliferation
XX CC and/or differentiation of stem cells. The ligands can be used, e.g.
XX CC chemotherapy, or radiation.
XX SQ Sequence 5404 BP; 1411 A; 1297 C; 1423 G; 1273 T; 0 other;

Query Match 17.3%; Score 713; DB 17; Length 5404;
Best Local Similarity 54.5%; Pred. No. 1.8e-134;
Matches 1711; Conservative 0; Mismatches 1345; Indels 81; Gaps 11;
QY 464 aacagaaaggagcgcacatgtgggtgcccctgtgtgtccatcccccgccctcaatgtcacg 523
DB 634 aacaagaacaaactgtgtgtatccccctgcggaggggtatttcaaacctcaatgtgtct 693
QY 524 ctgccc-----tcgcaaaagctcgtgtgtgtgcccagcggcagagaggtgtgtgagat 577
DB 694 ctltgcgtcgtaggtatcccaaaaagagattgttccggatggaacagaaatttccctgggac 753
QY 578 gaccagcgggcagctcgtgtccacgcccactgtgcacgagtcgctgtaccctgcagtcg 637
DB 754 agcgagataggttacttctccccagttacatgacagctatgccggcagtgcttctgt 813
QY 638 gagaccacctggggagaccagagacttcttccaaaccccttctgtgtgacatcacaggc 697
DB 814 gagcaaatcaatgtatgaacccatcagtcctatcatgtacatagttgtgtgtgagga 873
QY 698 aacgagctctatgacatccagctgttggccaggaagtcgctgagcgtcgtgtgagggag 757
DB 874 tataggattatgctgtattctgagcccccccgcatgataatgagcctatctgccggagaa 933
QY 758 aagctgtctgaactgcaccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 817
DB 934 aaacttgccttaattgtacagcagagaacagagctcaatgtgtgtgtgtgtgtgtgtgtgt 993
QY 818 gactaccacaggggaagcagagc-----agagcggggttaagtggtgtgtgtgtgtgtgtgt 869
DB 994 cacttccacccctcaaaagtctcatatagaagattgttaaacccggatgtgtgaaccccttt 1053
QY 870 cccagcagacccacacagaactctcc-aggatccttaccatccacacagctcagccagcagc 928
DB 1054 cctgggagctgtggcgaagatgttttggacaccttgacaaatagaagagtgaccaaagat 1113
QY 929 gacttgggtcgt 988
DB 1114 gaccaaggggaatacacacctgt 1173

Qy 989 gaggtcattgcatgaaatcccttcacacgctcagtgagtgctcaaaaggaccatccctg 1048
Dy 1174 ttgtccgagttcacacaaagcctttattgtcttggttagtggtggtgaaatcttgggtg 1233
Qy 1049 gagcccgccgagagagcagctgggtgaagctcccggtgaagctggcaycgtaacccecg 1108
Dy 1234 gaagccca---cagtgggcagtcgaagtcggaatcccggtgaagtagtctcagttaccacgt 1290
Qy 1109 cccgagttccagtggttacaaggatggaagggcactgtccggcgccaca-----gt 1159
Dy 1291 cctgatcaaatggtacagaaatgggaagggccattgagtcaccaactacacaatgattgtt 1350
Qy 1160 ccacatgccctgtgtctcaagaggtgaacagagccagcagcagcactacacacctcgcc 1219
Dy 1351 ggcgatgaactcaaccaactggaaggtgactgaaagagatgcaggaataactacacggtcacc 1410
Qy 1220 ctgtggaactccgtgtgctgagggcgcaacatacgaacctgagctgggtggaatgtg 1279
Dy 1411 ctaccaaaccattcaatggagaacacagagccacatggtctctcgtgtgtgaatgtc 1470
Qy 1280 ccccccagatacatgagaagaggctcctcccca---gcatctactcgcgtcacagc 1336
Dy 1471 ccacccagatcgtggaagagccttgatctgcctatgagttcctaccagatcgggacc 1530
Qy 1337 gcgcagggccctcacctgcacggcctacgggtgcccctgcctcagcatccagtgccac 1396
Dy 1531 atcgagaacttgacatgcacagtgactacgcgaacacctccctgcaccacacagtggtac 1590
Qy 1397 tggcgccctggacacctgcaagatgttgccagcgtagctccggcgccgagcagcag 1456
Dy 1591 tggcagctagaagaagctgc-----tctacagaccggccaaa 1630
Qy 1457 caagacctcatgcacacgtgcgtgactgagggcggtgaccacgcagagtgccgtgaac 1516
Dy 1631 caagcccgtagtc-----ttgtaagaaatggagacacgtggaggattccagggggggaac 1686
Qy 1517 cccatcgagagctggacacctggaccgagtttgtaggggaaagaataagactgtgagc 1576
Dy 1687 aagatcgaagtcacaaaacaaatagctgctgattgaaaggaaaaaacaacactgtaagt 1746
Qy 1577 aagctgggtgacagaaatgcaaacgctgctgcaggtatgataagtggtgtctcacaag 1636
Dy 1747 acgctgggtcatcaagctgcaaacgctgcaggtgtgataaatgtagaagccatcaacaaa 1806
Qy 1637 tggcgccagagtgagcggtcactactctatgtgaccacacatcccgacgggtccacc 1696
Dy 1807 gcgggacgagagagaggtcactctctccatgtgacaggggtccctga-----a 1857
Qy 1697 atcgaatccaagccatccgagagagctactagagggccgagcggtgctcctcagctgcaaa 1756
Dy 1858 attactgtcaacctgctgccagccactgagcaggagtggtgtcctctgtgtgact 1917
Qy 1757 gccacagctacaagtacagacatctgcgtggtaccgctcacaacctgtccacgtgcac 1816
Dy 1918 ggaagacaaatagctttgagaaacctacgtggtacaaagcttggttccctgtgtgact 1977
Qy 1817 gatgcgcaggaacccgctgtctgctgactgcaagaaagctgcatctgttcccccct 1876
Dy 1978 gtccacatggcgaaactcactacacagcttgcaagaaattgagctcttggaaactg 2037
Qy 1877 ctggccagcagctggagagagtggaacctggggcgccacgcacgcctcagctcagct 1936
Dy 2038 aatggcaca-----tgtttctaacagcacaaaatgacatcttgatttgca 2085
Qy 1937 atcccccgctgcgcgcgagcagcagggccactatgtgtgcgaagtgcaagaccggcg 1996
Dy 2086 ttcaagaaatgctctctgcaggaaccaagcgactatgttgcctgcacaaataaagaag 2145
Qy 1997 agccatgacaagcactgcacaaagaagtacctgtctggtgagggccctggaagccctcgg 2056
Dy 2146 accaagaaagacattgctgtgtaaacagctcatctactagagcgctggtcacccatg 2205
Qy 2057 ctacgcagaacttgaccacacctcctgtggaacgtgagcgactcgtggagatgcagtg 2116

Dy 2206 atcaccgaaatctggagaatcagacaacaccattggcgagaccattgaaagtactgtc 2265
Qy 2117 ttggtggccgagcgacgcgcgcacacgtgtggtacaaagacgagggctgctggag 2176
Dy 2266 cagcatctggaaatctaccocacacattacatgggttcaagacacagagacctggta 2325
Qy 2177 gaaaagtcggagtcgactggcgactccaacagaagctgagcatccagcgctgcgc 2236
Dy 2326 gaagattccagcatctactgagagatggaaacggaaacctgactatccgaggtgag 2385
Qy 2237 gaggaagtcgggacgcctatctgtcagcgtgtgcaacgcgaagggctgcactcc 2296
Dy 2386 aaggagtagaggtcctacacctccagggcgtgaatgtcctgtggtgctgcaagag 2445
Qy 2297 tccgcagcgtgcccgtggaagctccgaggaataaaggcagatggagatcgtgacct 2356
Dy 2446 gagacgtctcatalaatalaagaggtgcccaagaaagacaaactggaagtcattacc 2505
Qy 2357 gtcgggtaccggcgtcgtgtcttcttctgggtcctcctcctcctcctctctctgtaac 2416
Dy 2506 gtcggcactcagtgattgccatgtctcttctggtcctccttctcattctcgtacggac 2565
Qy 2417 atgaagagcgcccgccacgcagacatcaacagggctacgtgtccatcatcagacccc 2476
Dy 2566 gttaaagcgccaatgaagggaactgaagacagggctactgtctatgtgcatggatcca 2625
Qy 2477 ggggaggtgctcctggagggagcaatgcgaatcctgtcctacgatgccagccagtgaaa 2536
Dy 2626 gatgaattgccccttggaagcgtgtgaacgtgtgcttatgatgccagcaagtgga 2685
Qy 2537 tcccccgagagcgctgacacctggggagagtgctcgtgcagggcctcgggaaagt 2596
Dy 2686 tccccagggacggctgaaactaggaaaacctcttgccgcggtgctctcgcgcaagt 2745
Qy 2597 gtggaagcctccgtcttcgcacatccacaaggcgagcagctgtgacacgtggcgtgaaa 2656
Dy 2746 attgagcagacgctttggaattgacaagacagcagctgcgaacagtagcgtcgaag 2805
Qy 2657 atgctgaagagggcgccagcgccagcagcagcgcgtgctgctgcggagctcaagatc 2716
Dy 2806 atgtgaaagaagagcaacacacagcagcagcagcagcctcagtctgaactcaagatc 2865
Qy 2717 ctactacatccgcaaccacctcaacgtgtgtaacctcctcggggcggtgacacaaagc 2776
Dy 2866 ctcatccatgtgtccacctcctcaatgtggtgaacctctgagcctgagcgtgaccaaagc 2925
Qy 2777 caggccctcctatggtgagctgagttctgcaagtacgcaacacctctcacaactctctg 2836
Dy 2926 ggaaggccctcctatggtgtggaattctcgaattctcgaagtttgaaacctatcaactta 2985
Qy 2837 cgcgccaaagcgccctcagccctcagccctgcgggagagtgctcctccgagcagcggaagc 2896
Dy 2986 cggggcaagagaaatgaattgttccctataagagagaaaggcgacgtctccgcagggc 3045
Qy 2897 ttcgcgcctatgggtgagcctccagcagctgagtcggagggcgccgggagcagcgag 2956
Dy 3046 aagactacgttggtggagcctcctcgtgagctgaaaagacgttggacagcatcaccagc 3105
Qy 2957 gtcctcttcgcgcggttctcgaagacggcgagcgagcggggcttctccagaccaa 3016
Dy 3106 agccaagactcgtccagctcaggtctgttgagagaaatcgtcagtgatgtagaggaa 3165
Qy 3017 gaag-----ctgaggacctgtggtgagccgctgacacatgggaagatctgtctgctac 3070
Dy 3166 gaagaagctctggaagaaactgacaaggactcctgacttggagcactctcatctgttac 3225
Qy 3071 agcttccagctggccagagagtgagttcctggcttcccgaagatgcataccagagac 3130
Dy 3226 agcttcccaagtggtcgaagcaggtgagttcttggaatacaagaaagtgatccacagagac 3285
Qy 3131 ctgggctgctggaaactctgctgctcggaagcgagctggtggaagatcgtgactttggc 3190

Db	2386	aaggaggatggagacctatacactgccaggcctgcaagtgccttggctgtgcaagagcg	2444
Qy	2297	tccgcagcgttgccgtggaaagctccgaggataaggggcagcatggagatcgtgacctt	2356
Db	2446	gagacgctctcataatagaggtgcccgaaagacaaacttggagctattatcctc	2505
Qy	2257	gtcgtatccggcgatcgtctcttcttgggtgttccctccctcatctctgtaac	2416
Db	2506	gtcggcactgcagtgattgccattctcttgccctctcttctcatctctgtacggacc	2565
Qy	2417	atgaggagccggccacgcagacatcaagacgggtctacctgtccatcatctggacccc	2476
Db	2566	gttaagcgggccaatgaagggaactgaagacagcgtacttgtctattgtcatgatcca	2625
Qy	2477	ggggaggtgcctctgagagagcaatgcgaataactgtcctacgatgccagccagtgga	2536
Db	2626	gatgaattgcccttgatgagcgtgtgaecgttgcttatgtatgacgacagtgga	2685
Qy	2537	ttccccagagcgcgtgcacctggggagagtgcctgcgtacgcgccttcgggaagtg	2596
Db	2686	ttccccagggaccgcctgaaactagaaaaacctcttgccgcgtgtccttcggccaagt	2745
Qy	2597	gtggaagctccgcttttcggcatccacaaaggcagcagctgtgacacgtggccgtgaa	2656
Db	2746	attgaggcagacgtcttggaattgacaagacagcgacttgcaaaacagtagccgtca	2805
Qy	2657	atgctaaauagggcccgccagccagcagcacgcgcgtgatgtcggaagctcaagtc	2716
Db	2806	atgttgaagaagagcaacacagcggagctgagccctcatgtctgaaactcaagtc	2865
Qy	2717	ctcattcacatcggcaaccacctcaacgttggtcaacctcttcggggcgtgcaccaagccg	2776
Db	2866	ctcatccaacttggtcaecatccaatgtggtgaaacctctaggcgcgtgcaccaagccg	2925
Qy	2777	caggggccctctatgtgatgtgagttctgtcaagtagcggcagcccttccaaactctcg	2836
Db	2926	ggaggccctctatgtgatgtggaattctcgaagtttggaacctatacaactta	2985
Qy	2837	cgcgcacaagcgagcctctcagccctcgcgcggagaggtctcccagcagcgcgcgc	2896
Db	2986	cggggtcaagagaatuaattgttccctataagacgaaggggcacgctctccgcagggc	3045
Qy	2897	ttccgcgcctatggagagctccagagctggatctgagggcgcgcgggagcagcagc	2956
Db	3046	aaggactacgttggggagctctccgtggatctgaaaagacgcttgacagcatcaccagc	3105
Qy	2957	gtcctcttcgcgcggtcttcgaagaccagcgagcggaagcgggctctccagacaa	3016
Db	3106	agccagagctctccagctcagggcttggttgagagagaatcgtctcagtgatggatgg	3165
Qy	3017	gaag-----ctgaggaacctgtggctgagccgcgtgacctggaagatcttgtctgtac	3070
Db	3166	gaagaagcttctgaagaactgtacaaggaactcctgaccttggagcatctcatctgtac	3225
Qy	3071	agcttcagatggccagaggaatggaattccttggttcccgaatgcatccacagagac	3130
Db	3226	agcttcgaagtgtcaaggcattggagttcttggcatcaagaaagtgtatccacagagac	3285
Qy	3131	ctgggctctcggaacattctgctgcggaagcgcagctggtgaaagatctgtgacttgcc	3190
Db	3286	ctggcagcacgaacattctcctatccgagaagaatgtggttaagatctgtgacttcggc	3345
Qy	3191	cttgcgggacatctacaagaacctgactacgtccgcaaggcgagtgcctcgctgcc	3250
Db	3346	ttggcccggaacattataagaccgggatattgtcagaagaaggagatgcccgactccct	3405
Qy	3251	ctgaagtggatggccctgaagacatcttcgacaaggtgtacaccacagatgacgtgt	3310
Db	3406	ttgaagtggatggcccggaacacctttttgacagagtatacaacaattcagagcgatgtg	3465
Qy	3311	tggctcttggggtgctctctctggggagatctctctctctgggggctccccgtaccc	3370
Db	3466	tggcttctcagatgtgtgctctgggaatatttctcttaggtgctctcccatacccttgg	3525

[illegible]

Db	3286	ctggcagcagcaaacattctctatctcgagagaagaatgtygttaagatactgtgactctggc	3343
Qy	3191	ctlgccggagacatctacaaagacctgactacgtcccgcaaggcgagtgcccggctgccc	3250
Db	3346	ttgcccggagacatttataaaagcccgattatgtcagaaaaagagatgccgactccct	3405
Qy	3251	ctgaatgtagtggccctgaaagcattctcgacaaggtgtacacacgcagagtgacgtg	3310
Db	3406	tlgaagtggatggccccggaaaccatttttgacagagtatacacaaatlcagagcgtg	3465
Qy	3311	tgtctcttgggggtgcttctctggagagattctctctgggggctccccgtaccctggg	3370
Db	3466	tgtgtcttgggtgtgtgtctctggaaataattttctctaggtgtctccccataccctggg	3525
Qy	3371	gtgcagatcaatgagaggtcttgccagcggctgagagacggtgagacgcgcacaaagatgagggccccc	3430
Db	3526	gtcaagattgatgaagaattttgtaggagattgaaagaagaaactagaatgcgggtcct	3585
Qy	3431	gagctggcgaactcccgccataagccgcacatcatgtgaactgtgtcccgagagacccaag	3490
Db	3586	gactacactaccacagaatgtaccagaccatgtctggactgtgtgcatgagggacccaac	3645
Qy	3491	gcgagacctcattctcgagctggttgagagatcctcgggggacctctccaggcgaggggc	3550
Db	3646	cagagaccctggttttcagagttggttgagcatttgggaacctctgcgaagcaaatgcg	3705
Qy	3551	ctgcaagaaggaagagga	3567
Db	3706	cagcagatgcaaaaga	3722
RESULT 12			
Q81014			
ID	Q81014 standard; cDNA; 5406 BP.		
AC	Q81014;		
DT	18-AUG-1995 (first entry)		
DE	Fkl1 receptor protein-tyrosine-kinase cDNA.		
XX	Mouse Fkl1; receptor protein-tyrosine-kinase; hematopoietic stem		
KW	cell; fetal liver kinase; ds.		
OS	Mus musculus.		
FX	Key	Location/Qualifiers	
FT	CDS	208..4311	
FT		/*tag= a	
FT		/product= Fkl1 receptor protein-tyrosine-kinase	
FT	sig_peptide	208..264	
FT		/*tag= b	
FT	mat_peptide	265..4308	
FT		/*tag= c	
XX	WO9500554-A.		
XX	05-JAN-1995.		
XX	17-JUN-1994;	94WO-US06944.	
XX	18-JUN-1993;	93US-0080244.	
PR	21-JUN-1993;	93US-0081508.	
PR	23-NOV-1993;	93US-0157490.	
XX	(UYPR-) UNIV PRINCETON.		
XX	Lemischka IK;		
XX	WPI; 1995-052014/07.		
DR	P-PSDB; R67817.		
XX	Ligand for receptor protein tyrosine kinase - useful for the		

stimulation of primitive haematopoietic stem cells causing proliferation and/or differentiation

proliferation and/or differentiation

Disclosure: Fig 2: 131pp: English.

CC The sequence corresponds to a cDNA encoding a mouse Flkl (fetal
CC liver kinase) receptor protein-tyrosine-kinase. Flkl is expressed
CC in both primitive and mature hematopoietic cells, and in other
CC tissues. The gene product is useful in isolation of receptor
CC ligands, which have applications in diagnosis of bone marrow
CC disorders and in stimulating proliferation and/or differentiation
CC of primitive hematopoietic stem cells.

XX
C-700000 5406 PD. 1412 A. 1298 C. 1423 G. 1273 T; 0 other;

Query Match 17.38; Score 713; DB 16; Length 5406;

Query Match

Best Local Similarity	34.36;	Fixed NO: 1.00	1347
Mismatches	0;	Mismatches	1345;
Indels	81;	Indels	81;
Caps	11;	Caps	11;

523

[illegible]

693

Db 634 aacaagaacaaaactgtggtgatccccctgccgagggccgaalcccaaaaccccaatgctgccc

[illegible]

QY
524 ctgcgc-----tcgcaagctcgggtgctgtggccagacgggcaggagggc
|||||

753

Db 694 ctttgcgctaggtatccagaaaagagatttgctccggatcggaaacacagaaaccccccctgggac

-----tccacgaatcccctatacctcagtcg 637

[illegible]

813

Db 754 agcgagataggcttactctctccccagttacatgatcagctatgccggcatggtcccccgc

cacatcacagc

QY 638 gagaccctggggagaccaggacttcctttcccaaccccttccttggtgagacacacaggc

873

Db 814 gagcaagatcaatgatgaaacctatcagtctatcatgtacatgcccgcaggccggcgccg

[illegible]

QY 698 aacgagctctatgacatccagctgttgccagggaagtcgctggagctgctgctggagggag

933

Db 874 tataggatttatgatgtgattctgagccccccgcgcattgaaattgagctatctcgcggagaa 33

817

Qy 758 aagctggtcctgaactqacaccgtgtgggctgaggttttaactcagggtgtccacclllyaccl99 01

[illegible]

Db 934 aaactgtcttaaattgtacagcgagacagagctcaatgtgggcttgatttcacccctg

[illegible]

Qy 818 gactaccagggaagcaggc-----agagcgggtaagtgggtgcccgagcgacgct 863

Db 2686 tccccaggaccggtgaaactagaaaaacctcttgccgcggcggtgaccttcgcccgaagtg 2745
Qy 2597 gtggaagctccgtcttcggcatcccaagggcagcagctgtgaaacccgtgcccgtgaaa 2656
Db 2746 attgaggcagacgtcttggaaatgcaagacagcacttgcaaacagtagcgcgtcaag 2805
Qy 2657 atgctgaaagagcccgccagccagcagcagcgcgcgtgctgagctcgaagtc 2716
Db 2806 attgtgaaagagagcaacacacagcagcagcagcagcagcagcagcagcagcagcagc 2865
Qy 2717 ctattcacatcggaacacccctcaacgtggtcaacctctcggggtgagcagcagcagc 2776
Db 2866 ctcatccattgtgacacatcaatgtgtgaaacctctcagcgcctgcaacaaagccg 2925
Qy 2777 caggagccctctggtgagctggaattctgcaagtacggcaacctctcgaacttctg 2836
Db 2926 ggaaggccctctggtgagctggaattctgcaagtacggcaacctctcgaacttctg 2985
Qy 2837 cgcgccaaagcggagcgcctctcagccctgcgcgagagagctctcccgagcagcgcgc 2896
Db 2986 cgggcaagagaaatgaattgttccctataagagcaaaagggcagcgtctccgcagggc 3045
Qy 2897 ttcgcgcagctggtgagcagcagcagcagcagcagcagcagcagcagcagcagcag 2956
Db 3046 aaggactacgttgaggagctcctggtgagctgaaagacagcgttggaagcagcagcagc 3105
Qy 2957 gtccctctgcgcggttctcgaagacgcagcagcagcagcagcagcagcagcagcagc 3016
Db 3106 agcagagcctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 3165
Qy 3017 gaag-----ctgaggacactgtgctgagccgctgaccatggaagatcttctgtctac 3070
Db 3166 gaagaagctctgaagaactgtacaagacttctgaccttgaggagcagcagcagcagc 3225
Qy 3071 agcttccagctgagcag 3130
Db 3226 agcttccag 3285
Qy 3131 ctggtctgctggaacattctgctcggaagcagcagcagcagcagcagcagcagcagcag 3190
Db 3286 ctggcagcagcaaacattctctcagcagagagagagagagagagagagagagagagag 3345
Qy 3191 ctgcccagagacatcaaaacccctgactacgctccgcaagggcagctgcccggctgccc 3250
Db 3346 ttggcccgagacattataagagcccgagattatgtcagaagagagagagagagagag 3405
Qy 3251 ctgaagtgatggccctggaagcattcttcgaaggtgtacacacagcagagagagagag 3310
Db 3406 ttgaagtgatggcccggaacacattttgacagagatatacacacattcagagcagatg 3465
Qy 3311 tggctcttgggtgctctctcggagagatctctctggtggggtccctccctacccctggg 3370
Db 3466 tggctcttgggtgctctcggagagatctctctggtggggtccctccctacccctggg 3525
Qy 3371 gtgagatcaatgagaggtcttcgagcggctgagagagcggcagagagatgagggcccg 3430
Db 3526 gtcaagatgatgaagaatttggagagattgaaagagagagagagagagagagagagag 3585
Qy 3431 gaggctggccactcccgccatcagcgcagcagcagcagcagcagcagcagcagcagcag 3490
Db 3586 gactacactaccacagaaatgaccagagacatgctggagctgctggcagagagagagag 3645
Qy 3491 gcgagacctgactctcgagagctggtggagatcctgggggagacctgctccagggcgagg 3550
Db 3646 cagagacctcttctcagagagtggtggagcatttgggagacatttgggaaacctctc 3705
Qy 3551 ctgcaagagagagaga 3567
Db 3706 cagcaggatggcaaga 3722

RESULT 14
T72119

T72119 standard: cDNA: 5406 BP.
T72119:
19-AUG-1997 (first entry)
Murine flk-1 receptor coding sequence.
Human: fetal liver kinase 2; flk2; receptor protein tyrosine kinase;
p1K; liver; spleen; thymus; adult; brain; bone marrow/primitive portion;
haematopoietic hierarchy; extracellular domain; soluble form; ligand;
proliferation; differentiation; mammalian; haematopoietic stem cell;
macrocyclic anaemia; aplastic anaemia; cancer; radiation; mouse; ss.
Mus musculus.
Key Location/Qualifiers
CDS 208..4311
FT /*tag- a
FT /product= Murine_flk-1
XX US5621090-A.
XX 15-APR-1997.
XX 02-APR-1991; 91US-0679666.
XX 26-JUN-1992; 92US-0506397.
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX (UYP-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI: 1997-235228/21.
XX P-PSDB: W19875.
XX Protein containing the extracellular domain of human flk-2 - used
XX for identification of primitive haematopoietic cell proliferation
XX and differentiation stimulatory ligands, e.g. for treating anaemia
XX Disclosure: Fig 2: 55pp: English.
XX This sequence encodes the murine fetal liver kinase 1 (flk-1). flk-1
XX is a receptor protein tyrosine kinase (pTK). flk-1 is found in the
XX same organs as flk-2 (see also T72117-18) as well as in fetal brain,
XX stomach, kidney, lung, heart and intestine, and in adult kidney, heart,
XX spleen, lung, muscle and lymph nodes. flk-1 is not a member of the
XX same class of receptors as flk-2, as flk-1 may be found in more
XX mature haematopoietic cells. The invention concerns a recombinant
XX nucleic acid, preferably mRNA, which encodes a protein containing only
XX the extracellular domain of human flk-2 and lacking the flk-2 intra-
XX cellular catalytic domain. The resultant protein represents a soluble
XX form of flk-2 which is used to isolate specific ligands for flk-2. These
XX ligands can be used to stimulate proliferation and/or differentiation of
XX mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for
XX treatment of macrocytic or aplastic anaemia or bone marrow damage caused
XX by cancer treatment or radiation.
XX Sequence 5406 BP: 1412 A: 1298 C: 1423 G: 1273 T: 0 other;

Query Match 17.3%; Score 713; DB 18; Length 5406;
Best Local Similarity 54.5%; Pred. NO. 1.8e-134;
Matches 1711; Conservative 0; Mismatches 1345; Indels 81; Gaps 11;
Qy 464 aacaggaagcagccatgtggtgcccctgtggtgctccatccccggcctcaatgtcacg 523
Db 634 aacaagaacaaactgtgtggtgccccctgcccaggggtcgtttcaaacacccatgtgtct 693

Db 1012 aaactgtcttaattgtacagcgagacacagagctcaatgtggggtgtgatttcaactgc 1071
Qy 818 gactaccaggaagcaggc-----agagcgggtgaagtggtggtcccgagcgagcgt 869
Db 1072 cactctcaccctcaaaagctctcatcaagaagattgaaaccgggatgtaaacaccttt 1131
Qy 870 ccagcagacacacacagaaactctcc-agcactcctgaccatcccaacacgctagccagcac 928
Db 1132 cctgggactgtgcggaagatgttttgagcaaccttgacaalagaagtgtagcaagaagt 1191
Qy 929 gacctgggctcgtatgtgtcaagggcccaaacacggccatccagcgaatttcgggagagcacc 986
Db 1192 gaccaaggggaatacacctgtgtagctccagtggcggatgatcaagagaaatagaaca 1251
Qy 989 gaggctcattgtgcagaaatcccttcacagcgtcagtgagctcaagagaccatcctg 1048
Db 1252 ttgtccagagttcacaaagcccttlatgtcttcggtagtggaatggaalcttgggtg 1311
Qy 1049 gaggccacggcgagagcagctggtggaagctgcccgtggaagctggcagcgtacccccg 1108
Db 1312 gaagcca---cagtgggcagtcgaagatcgaaatccctgtgaagtatctcagttaccagct 1368
Qy 1109 cccgagttccagtgtgtacaaagatgaaagcactgtccggcgccaca-----gt 1159
Db 1369 cctgatcaaatggtacagaatggaagggcccatctgagtcacactacacaaatgatgtt 1428
Qy 1160 ccacatgccctggtgtcgaagagtgacagagccagacagcagccacacacacccctgccc 1219
Db 1429 ggcgtgaactcaccatcatggaagtgaactgaaagagatgacgagaaatcacagglcatc 1488
Qy 1220 ctgtggaactcgcgtgctggcctgagcgcgcaacatcagcctggagcgtggtggaatgtg 1279
Db 1489 ctcaaccaacccatttcaatgagaaacagagccacatggtctctctggttgtgaaatg 1548
Qy 1280 cccccagatcacatgagaaaggagccctccctcccca---gcactactcgtgtcacagc 1336
Db 1549 ccacccagatcgtgtgagaaagcctgtatctgcctatggaattcctaccagtatgggacc 1608
Qy 1337 gcacggccctcactgtcagcgcctacaggggtgcccctgcctctcagatccagtgagcac 1396
Db 1609 atgcacacattgacatgcacagtctacgcaaacccctcccttgcaacacatccagtggtac 1668
Qy 1397 tggcgccctggacacccctgcgaagtgttggccagcgtagcttcctcgcgcgcgagcag 1456
Db 1669 tggcagctgaagaagcctgc-----ccctacagaccggccaaa 1708
Qy 1457 caagacctcatgccacagtgccgtgactgagggcggtgacacgcaggatgcgtgaaac 1516
Db 1709 caagcccgatgc-----ttgtaagaatggagacacgtggagattccaggggggaaac 1764
Qy 1517 cccatcgagagcctggacacccctggaccaggttgttgaggggaaagaaataagactgtgagc 1576
Db 1765 aagatcgaagtacacaaaacaaatagcctgattgaaagaaacaaacaaactgtaagt 1824
Qy 1577 aagctggtgatccagaatgccaaactgtctgcgtatgacaaagtgtgtgtgtctccaaacag 1636
Db 1825 acgctgtcatcgaagctccaaactgtaagcgtgtgacaaatgtaagaaacaaactgtaac 1884
Qy 1637 gtgggcccagatgagcgcctcactctctatgtgaacacacaccccgagcgttccacc 1696
Db 1885 gcgggacgaggaagaggggtcatctctccatgctgtagcaggggtcctga-----a 1935
Qy 1697 atcgaatccaagccatcccgagagagctactagagggccagccgggtcctctgagctgcca 1756
Db 1936 attactgtcaacctgctgccagcaactgagcagagagtggtccctgtgtgtgact 1995
Qy 1757 gccacagctacagtaacagatctgcgtgtgaccgctcaaacctgtccacgctgcac 1816
Db 1996 gcagacagaataatcgttttgaaaccttcacgtgtgaaacagctgtggtcagagcaacatcg 2055
Qy 1817 gatcgccagggaaacccgcttctgtcagactgcaagaacagtgcatcgtttccgcccct 1876
Db 2056 gtccacatggcgaaatcaactcacaccagtttgcaagaacttggatgctctcttggaaactg 2115

Qy 1877 ctggcgccagcctggaggggtggcacctggggcgccagcgccacgcctcagcctgagt 1936
Db 2116 aatggacca-----tgtttctaacagacaaatgacatcttgattgtggca 2163
Qy 1937 atcccccggtgcgcgcggagcagcaggggccactatgtgtggaagtgcgaagaccggcgc 1996
Db 2164 ttcaagaatgcctctctgcaggaccaagcgactatgtttgtctctctcaagaataagaag 2223
Qy 1997 agcatgacaagcacgtgccacaagaagtacctgtcgggtgcaggccctggaaagccctcgg 2056
Db 2224 accaagaaagacattgcctgtgtcaaacagctcatcactctagagcgcgtggcaccatg 2283
Qy 2057 ctcaacagaaactgaccgacctcctcctggtgaaactgagcagcctgcgtgagatgcagtcg 2116
Db 2284 atcacggaatctggagaalacagacaacacccattggcgagaccattgaagtgaacttcg 2343
Qy 2117 ttggtggccgagcgcacgcgccagcatcgtgtgttacaagaacgagagcgtgtggag 2176
Db 2344 ccagcatctggaatcctacccacacattacatggttcaagaacacagagaccctggta 2403
Qy 2177 gaaaagtctgagtcgacttggcgactcccaacagaagctgagcatccagcgcgtgcgc 2236
Db 2404 gaagattcaggcattgtactgagagatgggaacgggaacctgactatcccgaggtgag 2463
Qy 2237 gaggagatgcgggacgtctatctgtcagcgtgtgcaacgccaagggcgtcgtcaactcc 2296
Db 2464 agggagatgagggcctctacacctgcaggcctgcaatgcttggctgtgcaagagcg 2523
Qy 2297 tcccgccagcgtggcggtggagcgtccgaggtataaaggcgagcatgagatcgtgactct 2356
Db 2524 gagacgtcttctcaatagaaagtgcccgaggaagacaaacttggaaagtcattatctc 2583
Qy 2357 gtcagtcacggcgctcagctgtctcttctgtggtcctcctcctcctcctctctgtaac 2416
Db 2584 gtcggcactcagtgatgctcctgtctgtcctcctcctcctcctcctcctcctcagacc 2643
Qy 2417 atggaggggccggcccgagacatcaagacgggctacgttccatcctacatcagacccc 2476
Db 2644 gtaagcgggccaatgaagggaactgaacacaggtactgtctatgtgcatggatcca 2703
Qy 2477 ggggagtgtcctctggaggagcaatgcgaatatacctgtcctacgatgccagccagtgaaa 2536
Db 2704 gatgaattgccttggatgagcgtgtgaacgtctcctctatgatgcagcaagtgaaa 2763
Qy 2537 tccccgagagcgctgcacccctgggagagtgctcggtacggcgcttcctcggaagagtg 2596
Db 2764 tccccagggaccgctgaaactagaaaacctcttggcgcggtcccttcggccaagt 2823
Qy 2597 gtggaagcctccgcttctggcatcccaagggcgagcagctgtgacacgtggccgtgaaa 2656
Db 2824 attgagcgagcgtcttggaaatgacaagacagcagcttgcacaaacagtagcgtcaag 2883
Qy 2657 atgctgaaagaggcgccagccagcgacgcgcgctgagtgctggagctcaagatc 2716
Db 2884 atgttgaagaagagagcaacacagcgagcatcgacctcatgtctgaaactcaagatc 2943
Qy 2717 ctcatcactcggcaacccctcaacgtgttcaacctcctcctcggggcggtgcacaaagccg 2776
Db 2944 ctcatccacattggtcacacctctcaatgtggtgaacctcttagcgctgcacaaagccg 3003
Qy 2777 caggccccctcagtgatcgttgagttctgcaagtacggcaacccctcctcctcctcctg 2836
Db 3004 gggagggcctcctcagtggtgtggtgaaatctgcaagtttgaaacctatcaactactta 3063
Qy 2837 gcgcgcaaggggacgccttcagccctgcgcgagaaagtctcccgagcagcgcggaagcgc 2896
Db 3064 cggggcaagaaatgaaattgttccctataagagcaaggggcagcgttcctcgcaagggc 3123
Qy 2897 ttccgcgcacatggtgagcctcgcagcgttgatcggagggcgggggagcagcgacag 2956
Db 3124 aaggactacgttggggagcctcctcgtggtatctgaaagacgcttggacagcatcaccagc 3183


```
Db 3541 GGGCAGGGCCCTGCAAGAGGAAGAGGAGTCTGTCATGTGCCCCCGGCGAGCTCTCAGAGCTC 3600
Qy 3601 agaagaggcagctctcgcagggtccacatggtccctacacatcgccaggtgacgc 3660
Db 3601 AGAAGAGGGCAGCTTCTCGCAGGTGTCCACCATGGCCCTACACATCGCCCCAGGCTGACGC 3660
Qy 3661 tgagagacgcccgaagcctgacgagccacagcctggtgcccaggtattacaactgggt 3720
Db 3661 TGAGGACAGCCCGCAAGCCTGACGCGCCACAGCCTGGCCCGCCAGGTATTACAACCTGGGT 3720
Qy 3721 gtccttccggggtcctggccagaggggtgagaccgtggttccctccagagatgaagac 3780
Db 3721 GTCTTTTCCCGGGTCCCTGGCCGAGAGGGGCTGAGACCCGTGTTCTCCAGGATGAAGAC 3780
Qy 3781 atttgaggaattcccatgaccccccaacacctacaaggctctgtggacaaccagacaga 3840
Db 3781 ATTTGAGGAATTCCTCATGACCCCAACGACCTACAAAGGCTCTGTGGACACACGACAGA 3840
Qy 3841 cagtggaattggtgctggcctggagaggtttgagcagatagagagcagagcatagaaga 3900
Db 3841 CAGTGGGATCGTCTGGCCTCGAGAGGAGTTTGAGCAGATAGAGACGAGCATAGACAAGA 3900
Qy 3901 aagcgctcagctgtaagagacctggccagaaatgtgctgtgacaggggcacacccctga 3960
Db 3901 AAGCGCTTCAGCTGTAAAGGACCTGGCCAGAAATGTGGCTGTGACAGGGCACACCTGA 3960
Qy 3961 cccccaaaggagggcgccgctgagcggggggcccgaggggcccgggtgttttcaaa 4020
Db 3961 CTCCCAAGGAGGCGCGCGGCTGAGCGGGGGCGCCGAGGAGGCCAGGTGTTTACAA 4020
Qy 4021 cagcagatggagagctgcagaccagagcagagagagagagcactgctcccgtctgccc 4080
Db 4021 CACCGAGTATGGGAGCTGTCCGAGCCAAAGCGAGGAGGACCACTGCTCCCGCTGTCGCCG 4080
Qy 4081 cgtgactttcttcacagacaacagctactaa 4111
Db 4081 CGTGACTTTCTTCACAGACAACAGCTACTAA 4111

RESULT 2
LOCUS AR016568 4195 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 1 from patent US 5776755.
ACCESSION AR016568
VERSION AR016568.1 GI:3972845
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4195)
AUTHORS Alitalo,K., Aprelikova,O., Pajusola,K., Armstrong,E., Korhonen,J.
and Kaipainen,A.
TITLE FLT4, a receptor tyrosine kinase
JOURNAL Patent: US 5776755-A-1-02-Jul-1998;
FEATURES
source location/Qualifiers
BASE COUNT 889 a 1279 c 1305 g 722 t
ORIGIN

Query Match 95.2%; Score 3912.4; DB 81; Length 4195;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3913; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ccacgcagcagccgagagatgacgagggcgccgctgtgctgcgactggtgctctg 60
Db 1 CCAGCGCAGCGCCGAGATGACGCGGGCGCGCGCTGTGCTCGCAGACTGTGGCTCTG 60
Qy 61 cctgggactcctgagcggcctggtgagtggttactccatgaccccccgacattgaacat 120
Db 61 CTTGGGACTCTCGAGCGGCTGTGTGAGTGCTACTCCATGACCCCCCGGAACTTGAACAT 120
```

```
Qy 121 cagcagagagtcacagctcatcgacaccgggtgacagcctgtccatctcctgcaggggaca 180
Db 121 CACCGAGAGCTCAGCTCATCGACACCGCTGACAGCGCTGCCATCTCCTGCGAGGGACA 180
Qy 181 gcacccctcagtggtggtcagcagggagctcaaggaggcgccacccggaagacaagga 240
Db 181 GCACCCCTCTGAGTGGGCTTGGGCGAGGAGTCAGGAGGCCGCCAGCCACCGGAGACAGGA 240
Qy 241 cagcagagacacgggggtggtgcgagactgcgaggggcacagacyccagggccctactgcaa 300
Db 241 CAGCGAGACACGGGGGTGTGCGAGACTGCGAGGGGCACAGACGCCAGCGCCCTACTGCAA 300
Qy 301 ggtgtgctgctgcacagagtlacatgccaaacgacacagggcagctacgtctgtactacaa 360
Db 301 GGTGTGCTGCTGCACAGGTATATGCCAAACGACACAGGCGAGCTAGCTGTGCTACTACAA 360
Qy 361 gtacatcaaggcagcagtcagggcaccacggcgcgcagctcctacgtgttctgtgagaga 420
Db 361 GTACATCAAGCGACGCTATCGAGGCGACACGCGCGCGAGCTCCTACGTGTTCGTGAGAGA 420
Qy 421 cttgagcagccattcatcaaacgctgcacgctcttggccaacaggaagagcgccat 480
Db 421 CTTTGAAGCAGCATTCATCAACAGCGCTGACACGCTTTGGTCAACAGAGGAGCGCCAT 480
Qy 481 gtgggtgcccctgtgctgtgctccatccccggcctcaatgtcacgctgcgctcgcaaaagtc 540
Db 481 GTGGGTGCCCCTGTGCTGTGCTCCATCCCCGGCCTCAATGTACGCTGCGCTCGCANAAGCTC 540
Qy 541 ggtgctgtggccagcagggcagggaggtggtgtggatgacccggcgggggcgtctgtgtc 600
Db 541 GGTGCTGTGGCCAGCAGCGGAGGAGGTGGTGTGGATGACCGCGCGGGCATGTCTGTGTGTC 600
Qy 601 cagcgcactgctgcacagatgacctgtacctgcagtgagacaccacctgggagacacagga 660
Db 601 CAGCCCACTGCTGCACAGATGCCCTGTACTCTGCAGTGGAGACCACTTGGGAGACCAAGGA 660
Qy 661 cttcctttccaaccccttccctggtgcacatcacaggcaacgagctctatgacatccagct 720
Db 661 CTTCTTTCCAACCCCTTCCTTGGTGACATCACAGGCAACGAGCTCTATGACATCCAGCT 720
Qy 721 gttccccaggaagtcgctgagagctgctggtaggggagaaagctggtcctgaactgcacgt 780
Db 721 GTTCCCGAGGAAGTCTGCTGGAGCTGCTGGTGGAGGAGAAAGCTGGTCTGAACTGCACCGT 780
Qy 781 gtgggctgagtttaactcaggtgtcacctttgactgggactccccagggagcagggcaga 840
Db 781 GTGGCTGAGTTTAACTCAGGTGTCACTTTGACTGGGACTACCCAGGAAAGCAGGAGA 840
Qy 841 gcggggtaagtgggtgctgagcagcagcctcccagcagacccacacagaactctccagcat 900
Db 841 GCGGGTAAAGTGGGTGCTGCGAGCGACCGCTCCACAGACCCACACAGAACTCTCCAGCAT 900
Qy 901 cctaacatccacaacagctcagccagcagcagcctgggctcgtatgtgagagggccaacaa 960
Db 901 CCTGAACATCCACAACGCTCAGCCAGCAGCAGCTGGGCTCGTATGTGTGCAAGGGCCAACA 960
Qy 961 cggcattccagcgatttcggggagagcagcggttcattgtgcatgaaatcccttcacag 1020
Db 961 CGGCATCCAGCGATTTCGGGAGAGCACCGAGGTCAATTGTGTCATGAAAATCCCTTCATCAG 1020
Qy 1021 cgtcagatggctcaaaagaccatcctggaggccacgagcagagcagcagctgggtgaagct 1080
Db 1021 CGTCAGTGGCTCAAAAGACCCATCTCTGGAGGCCACCGCAGCAGCAGCAGCTGTGTGAGCT 1080
Qy 1081 gcccgtagagctggcagcagctaccccccccgagttccagtggttacaaggtggaagggc 1140
Db 1081 GCCCGTAGAGCTGGCAGCGTACCCCCCGGCGAGTTCAGTGGTACAAGGATGAAAGGC 1140
Qy 1141 actgtccgggcgcacagtcacatgcccgtgtgctcaaggaggtgacagagggcagcac 1200
Db 1141 ACTGTCCGGGCGCCACAGTCCACATGCCCTGTGCTCAAGGAGGTGACAGAGGCCAGCAC 1200
```

QY 1201 aggcactacacctcgccctgtggaactccgctgctggcctgagcgcaacatcagcct 1260
|||||
Db 1201 AGGCACCTACACCTTGGCCCTGTGGAACCTCCGCTGCTGGCCTGAGCGCAACATCAGCCT 1260
QY 1261 ggaagctgggtggaatgtgtcccccagatacatagagaaggagggcctcctccccagatc 1320
|||||
Db 1261 GGAGCTGGTGGTGAATGTGCCGCCAGATACATAGAGAAGGAGGCTCTCTCCCCAGCAT 1320
QY 1321 ctactcgctcaacccgagccagccctcaactcgagcgcctacgaggtgccccctcctet 1380
|||||
Db 1321 CTACTCGCTGCAGCGCGGAGCCCTCACCTTCACCTGCACGGCTACGGGGTGGCCCTGCCCT 1380
QY 1381 cagcatccagtgcaactggcgccctggacacccctgcaagatgttttgocagctagctet 1440
|||||
Db 1381 CAGCATCCAGTGGCACTGGCGGCCCTGGACACCTTGCAAGATGTTTGGCCAGCTAGTCT 1440
QY 1441 ccggcgggcagcagcaagaacctcatgccacagtgccgtgactgagagggcagtgaccac 1500
|||||
Db 1441 CCGGCGGGCGCAGCAGCAAGACCTCATGGCCACAGTGCCGTGACTTGAGGGCGGTTGACCCAC 1500
QY 1501 gcaggatgcccgtgaacccccatcgagagcctggacacctggacccctggagcttttgtgaggaaa 1560
|||||
Db 1501 GCAGGATGCCGTGAACCCCATCGAGAGCCTGGACACCTGGACCCAGTTCGTGGAGGGAAA 1560
QY 1561 gaataagactgtgagcaagctgtgtatccagaatgccacgtgtctgccatgtacaaagt 1620
|||||
Db 1561 GAATAAGACTGTGAGCAAGCTGTGTATCCAGATGCCAAGTGTCTGCCATGTACAAAGTG 1620
QY 1621 tgtggtctccaaagtgggccagatgagcggtcctatctacttctatgtgaccacct 1680
|||||
Db 1621 TGTGCTCTCCAACAAGTGGCCAGGATGAGCGGCTCATCTACTTCTATGTGACCACTCAT 1680
QY 1681 ccccgagcgttcaactcgaatccaaagccatccgaggaagctactagagggccagccggt 1740
|||||
Db 1681 CCCCGAGCGGTTCAACCATCGAATCCAAAGCCATCCGAGGAGCTACTAGAGGCCAGCCGGT 1740
QY 1741 gctcctgagctgccaagccgacagctacaagtacagagcatctgcgtggtlaccgcctcaa 1800
|||||
Db 1741 GCTCCTGAGCTGCCAAGCCGACAGCTACAAGTACGAGCATCTCGCTGCTACCGCTCAA 1800
QY 1801 cctgtccacgctgcacgatgcacgggaacccgctctctgcctgcagctgcaagaacgtgca 1860
|||||
Db 1801 CCTGTCCACGCTGCACGATGCGACGGGAACCCGCTTCTGCTGACTGCAAGAAGCTGCA 1860
QY 1861 tctattccacacctctgcccgcagccttgagagaggtggaacctggggcgcgccacgc 1920
|||||
Db 1861 TCTGTTGCCACCCCTCTGGCCGCCAGCCTTGGAGGAGGTGGCACTTGGCGCGGCCACGC 1920
QY 1921 cagctcagcctgagtatcccccgctgctgcgcccagcagcagggccactatgtgtcga 1980
|||||
Db 1921 CAGCGTCAAGCTGAGTATCCCGCCGCTCGCGCCGAGCAGAGGGCCACTATGTGTGCGA 1980
QY 1981 agtgcaagacggcgacccatgcaagcactgccaagaagtacgtgtcggtgagggc 2040
|||||
Db 1981 AGTGAAGACCGCGCAGCCATGATGAAGCACTGCCAAGAACTACCTGTGTGTGAGGC 2040
QY 2041 cctggaagccccctggctcagcagaaactgaaccgacctcctggtgaacgtgagcactc 2100
|||||
Db 2041 CCTGGAAGCCCTCGGCTCAGCAGAACTTGACCGAGCTCTCTGTGTGAACGTGAGCGACT 2100
QY 2101 gctggagatgagtgctgtgtggcggaagcgacgcccagcatcgtlggtacaaga 2160
|||||
Db 2101 GCTGGAGATGAGTGTGTTGTTGGCGGAGCGCACGCCCCAGCATCTGTGTGTACAAGA 2160
QY 2161 cgaagagctgctggaggaagaaagtctgagtcgacttggggactcccaacagaagctgag 2220
|||||
Db 2161 CGAGAGGCTGTGGAGGAAAAGTCTGGAGTCTGACTTGGCGACTCCCAACCAAGAGCTGAG 2220
QY 2221 catccagcgctgctgcgagggaggtgcgggagcgtactctgtgcagcgtgtgcaagcga 2280
|||||
Db 2221 CATCCAGCGCTGCCGAGGAGGATGCGGGAGCGCTATCTGTGCACCGCTGTCGAACGCCAA 2280
QY 2281 gggctgctcaactcctccgcagcgctggccgtggaaggctccgaggaataaggcgagcat 2340
|||||

Db 2281 GGGCTGGCTCAACCTCCCGCAGCGTGGCCGTGGAAGGCTCCGAGGATGAAGGCAGCAT 2340
|||||
QY 2341 ggaatcgatgactctgtcggtacggcgctcatcgctgtcttcttctgggtctcctcct 2400
|||||
Db 2341 GGAGATCGTGATCTGTGCGGTACCGCGCTCATCGTGTCTTCTTCTGGGTCTCTCTCT 2400
QY 2401 cctcatctctgtaacatgagagggcgcccaagcagcaacatcaagaggggctacatctc 2460
|||||
Db 2401 CCTCATCTCTGTAACATGAGAGGCGCGGCCACGACATCAAGACGGGCTACCTGTCT 2460
QY 2461 catcatcagaccccgaggaggtgctcctgagaggaatgcaatacctgtcctacga 2520
|||||
Db 2461 CATCATCATGACCCCGGGAGGTGCCCTCTGGAGGAGCAATGCCANATACCTGTCTCTACGA 2520
QY 2521 tggcagccagtgggaattccccccgagagcggtgcacctgggggagtgctcggtcagcg 2580
|||||
Db 2521 TGCCAGCCAGTGGGAATTCGCCCGAGAGCGGCTGCACCTGGGGAGAGTGTCTCGGCTACGG 2580
QY 2581 cgccttcgggaaggtgtgtggaagcctccgcttcggtccatccacaagggcagcagctgta 2640
|||||
Db 2581 CGCCTTCGGGAAGGTGGTGGAAAGCCTCCGCTTTCGGCATCCCAAGGCGCAGCTGTGA 2640
QY 2641 caccgtggcgtgaaatgctgaaagagggcgccacggccagcgagcaccgcgcgctgat 2700
|||||
Db 2641 CACCGTGGCGGTGAAATGCTGAAAGAGGGCGCCACGGCCAGCAGCAGCAGCGCGCTGAT 2700
QY 2701 gtcggagctcaagatcctcattcattcagcctcggcaaccctcaacgtggttcaacctcctcg 2760
|||||
Db 2701 GTGCGAGCTCAAGATCTCTTTCATCTCGGCAACCACTCAACGTGTGTCAACCTCTCTCG 2760
QY 2761 ggcgtgacacagcgagggccccctcatggtgatcgtggagttcttcaagtagcagga 2820
|||||
Db 2761 GCGTGCACCAAGCCGACAGGCGCCCTCATGTGTGATCGTGGAGTTCGTCAAGTACGGCAA 2820
QY 2821 cctctccaaactcctgcgccgcaagcgagccttcagccctcgcgaggaagtctcc 2880
|||||
Db 2821 CTTCTCAACTTCTTCGCGGCCCAAGCGGACGCTTACGCCCTCGCGGAGAAAGTCTCC 2880
QY 2881 cgaagcgcggagcgttcgcgcgcatggtgagctcgccagcgtggatcgaagcgcc 2940
|||||
Db 2881 CGAGCAGCGCGGACGCTTCGCGCCATGTGTGGAGCTCGCCAGCTGGATCGGAGCGCC 2940
QY 2941 ggggagcagcagaggttctcttcgcggttctcgaagaccgagggcgagcgagggcg 3000
|||||
Db 2941 GGGGAGCAGCAGAGGCTCTCTTCGCGGCTTCTCGAAGACGAGGGCGGAGCGAGGGCG 3000
QY 3001 ggcctctccagaccaaagactgaggaacctgtggtgagccgctgacctggaagatct 3060
|||||
Db 3001 GGGTCTCTCCAGACCAAGAGCTGAGGACCTGTGGCTGAGCCGCTGACCATGGAAGATCT 3060
QY 3061 tgtctgacagcttccaggtggccagaggaatggagttctctggcttcccgaaagtgc 3120
|||||
Db 3061 TGTCTGTACAGCTCCAGGTGGCCAGAGGATGGAGTTCCTGGCTTCCCGAAAGTGCAT 3120
QY 3121 ccaagaacctgctgctcggaacctctgctcggaagcgagcggtggaagatctg 3180
|||||
Db 3121 CCACAGAGACTGCTGCTCGGAACATTCGCTGTGGAAGGCGACGTGGTGAAGATCTG 3180
QY 3181 tgaacttggccttggccgggacatctacaagacctgactacgctccgcaaggcgagtc 3240
|||||
Db 3181 TGACTTTGGCCTTGCCCGGGACATCTACAAGACCCCTGACTAGTCTCCGCAAGGGCAGTGC 3240
QY 3241 ccggctgccccctgaagtggatggccctgaaagcatcttcgaaggtgtacaccagca 3300
|||||
Db 3241 CCGGCTGCCCTGAAGTGGATGGCCCTGAAGACATCTTCGACAAAGGTGTACACAGCA 3300
QY 3301 gagtgaagctgtgttcttgggtgcttctctggagatcttctctgtgggggctcccc 3360
|||||
Db 3301 GAGTGAAGCTGTGTCTTTGGGTGCTTCTCTGGAGATCTTCTCTGTGGGGCTCTCCCC 3360
QY 3361 gtaccctgggggtgcagatcaatgaggagttcttgcacggctgagagcgagcaagat 3420
|||||

ORIGIN

[illegible]

and Birnbaum,D.
Chromosomal localization of FLT4, a novel receptor-type tyrosine kinase gene
Genomics 13 (2), 475-478 (1992)
JOURNAL 92307693
MEDLINE 3 (bases 1 to 4450)
REFERENCE Galland,F.
AUTHORS Direct Submission
TITLE Submitted (28-DEC-1992) F. Galland, INSERM, Unite 119, 27 Bd Lei
JOURNAL Roure, Marseille 13009, FRANCE
FEATURES
source Location/Qualifiers
1..4450
/organism="Homo sapiens"
/db_xref="taxon:9606"
22..87
sig_peptide
CD5
/function="tyrosine kinase"
/note="unnamed protein product"
/citation=[1]
/codon_start=1
/protein_id="CAA49505.1"
/db_xref="GI:297050"
/db_xref="SWISS-PROT:P35916"
/translation="MORGAALCLRLNLGLLDGLVSDYSMTPTLTNITEESVIDTG
DSLSTSCRGHLEWMPGAQAPATGDKDSEDGVVRCDEGTDARPYCKVLLVHVH
ANDGYSVYKVKIARIETGAASSYVEFDFEQPFINKPOTLLVNRKIDAMVPCLV
SIPGLNVTLSQSSVLVDPGQGVVMDRGMVSTPLLDHALYLOCEETWGDQDFLSN
PFLVHTIGNEDLDIQLPRKSLLELLVGEKLVNCTVWAEFNSGVTFDMYVPGQAEKG
KWPKRRSQOHTTELSILTTHNSQHDGLGYSVCKANNGIORPRESEVIVHENPFLS
VEMLPQTLEATAGDELVKLPVLAAYPPPEFQYKQKALSGRHSFHALVLEVEIA
STGYTTLAHNSAAGLRNRTSLELVNVPPOIHEKEASSPSIYSRHSRQALITCTAYGV
PLPLSIOWHWPMTCKMFAQRSLRRRQODLMPQCRDRAVTTQDAVNPIESLDWT
BEVECKNTVSKVLIONANVSAMKCVVSNKVGQDERLIYFYVTTIPDGFTEKPSLE
ELLEGCPVLLSQADSKYKHLRWYRLNLSLTHDAGNPLLLDCKNVHLPATPLAASL
EVAFAKHATLSLSIPRAPEHEGHVCEVQDRSHDKHCKKYLKSLQVLEAPRLTQ
NLTDLLVNSDLMQCLVAGAHAPSIVWYKDERLLEKSGVDLADSNOKLSIORVE
EDAGPYLCVSRPKGVNNSASVAVESGDSMEIVILVGTGVIAVEFVWVLLILFC
NNRPAHADIKTYLSTIMDPGEVEQCEYLSTIDASQWEPFRERHLGRVLGYGAF
GKVVESAFGIHKSSCDTAVKMLKEGATASEQALMSCLKILIHGNHILNVNLLG
ACTKPGQPLVIEFVKYGNLSFLRAKRDAPSCAEKSPQGRFRAMVELAKDRR
RPGSDRLVFAFSTEGCARRASPDQEAEDLMSPLTMDLVCSFOVARGMEFLAS
RKTCHDLAANLILLESDDVKICDFGLARDIYKDPDYVRKGSARLPLKWMAPESIFD
KYVTTQSDWFGVLLHFEISLGSAPYGVQINEEFCORVRDGRMRAPELATPAIRH
IMLNCWSDPKARPAFSOLVELLGLDGLQGLQEEVCHAPRSQSESEGSQVST
MALHTAQADAEDSPSLQRHSLAARYNNVSPFQCLARGAETRSSRKNKTFEFPMT
TYYKGSVNDTDSGMVLASEEPEQIESRRHROESGR"
22..2346
misc_feature
/citation=[1]
/product="extracellular domain"
misc_structure
160..366
/citation=[1]
/function="IG-like domain 1"
misc_structure
481..651
/citation=[1]
/function="IG-like domain 2"
misc_structure
763..963
/citation=[1]
/function="IG-like domain 3"
misc_structure
1072..1230
/citation=[1]
/function="IG-like domain 4"
misc_structure
1342..1635
/citation=[1]
/function="IG-like domain 5"
misc_structure
1741..1992
/citation=[1]
/function="IG-like domain 6"
misc_feature
2104..2277
/citation=[1]
/product="IG-like domain 7"
misc_feature
2347..2412
/citation=[1]
/citation=[2]

misc_feature
2413..2547
/product="transmembrane domain"
/citation=[1]
/citation=[2]
misc_feature
2548..2853
/product="juxtamembrane domain"
/citation=[1]
/citation=[2]
misc_feature
2854..3048
/product="tyrosine kinase domain 1"
/citation=[1]
/citation=[2]
misc_feature
3049..3558
/product="kinase insert"
/citation=[1]
/citation=[2]
misc_feature
3559..3786
/product="tyrosine kinase domain 2"
/citation=[1]
/citation=[2]
BASE COUNT 960 a 1350 c 1354 g 786 t
ORIGIN
Query Match 94.6% Score 3890; DB 78; Length 4450;
Best Local Similarity 99.6% Pred. No. 0;
Matches 3899; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 ccacgcgcagcggccggagagatcagcggggcgccgcgcgtgctcgcgacttggtctcgtg 60
DB 3 CCACGGCCACGGCCGGAGATGAGCGGGCGCGCGCTGTGCTCGGACTGTGGCTCTG 62
QY 61 cctgggactcctggacgcctggtgagtggtactcctcctgacccccccgacctgaaatc 120
DB 63 CTTGGGACCTGGAGCGCTGGTGTGAGTGACTACTTCATGACCCCGGACCTTGAACAT 122
QY 121 cagcgaggagtcacacgcctcgcacacgcctgagcgcctcctcctcgcagggagaca 180
DB 123 CACGGAGGAGTCACATGTCATCGACACCGGTGACAGCCTGTCCATCTCTCGAGGGGACA 182
QY 181 qcacccctcagtgggccttggccagagctcagcagggcgccagccaccggagacaaaga 240
DB 183 GCACCCCTCGAGTGGCTTGGCCAGGAGCTCAGGAGGCGCCAGCCAGGAGACAAGA 242
QY 241 cagcgagacacggggggtggtgagagactcgaaggcagcagcagcagccctactgcaa 300
DB 243 CAGCGAGACACGGGGGTGTGCGAGACTGCGAGGGCACAGACGCCAGGCCCTACTGCAA 302
QY 301 ggtgttgcctgcgcagaggtacatgccaacqacacagcagcagctacgtctgctactaca 360
DB 303 GGTGTGTGCTGTCAGCAGGTACATGCCACACACAGCAGCTAGGTCTGCTACTACAA 362
QY 361 gtacatcaaggcagcagcatcaggggcaccacgcgcgcgcgcctcctcgtcttcgtgagaga 420
DB 363 GTACATCAAGGCACGCATCGAGGGCACCCAGCGCCAGCTCCTACGTGTGCTGAGAGA 422
QY 421 ctttgcagcagcattcaacaacgcctcgcacgcctcttgcacaggaagggagcagccat 480
DB 423 CTTTGAGCAGCATTCATCAACAGCCCTGACACGCTCTTGTTGTCACAGGAGGAGCCAT 482
QY 481 gttgggtgcctcgtctgtgtcctcctcccgccctcaatgtcacgctgcctgcgaaagctc 540
DB 483 GTGGGTGCCCTGTCTGGTGTCCATCCCGGCCCTCAATGTACGCTGCGCTGCGAAGCTC 542
QY 541 ggtgctggtgcagcagcggcagggaggtggtggtggtggtggtggtggtggtggtggtg 600
DB 543 GGTGTGTGTCAGCGGGCAGGAGGTGTTGGGTATACCGCGGGGGGCGCATGCTGCTGTC 602
QY 601 cagcgcactgctgcagcatgcctcctcctcctcctcctcctcctcctcctcctcctcctc 660
DB 603 CACGGCCTGCTGCAGCATGCTGTTACCTGCAGTGCAGAGACCACTCTGGGGAGGAGCA 662
QY 661 cttcctttccaaacctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 720
DB 663 CTTCTTTTCAACCCCTTCTGTTGTCATCATCAGGCAACGAGCTCTATGACATCAGCT 722

Tue Jan 23 12:06:46 2001

Qy 721 gttcccccaggaagtcgctggagctgctggttaggggaaagactgctcgaactgcacccgt 780
 Db 723 gttcccccaggaagtcgctggagctgctggttaggggaaagactgctcgaactgcacccgt 782
 Qy 781 gtgggctgagtttaactcaagtgctacactttgactgggactaccagggagcagcaga 840
 Db 783 gtgggctgagtttaactcaagtgctacactttgactgggactaccagggagcagcaga 842
 Qy 841 gcggggtaagtgggtgcccagagcagcgtcccagcagcagcagcagcagcagcagcagc 900
 Db 843 gcggggtaagtgggtgcccagagcagcgtcccagcagcagcagcagcagcagcagcagc 902
 Qy 901 cctgaccatcacaaactcagccagcagcagcagcagcagcagcagcagcagcagcagc 960
 Db 903 cctgaccatcacaaactcagccagcagcagcagcagcagcagcagcagcagcagcagc 962
 Qy 961 cggcatccagcgaatttcggggagcagcagcagcagcagcagcagcagcagcagcagcag 1020
 Db 963 cggcatccagcgaatttcggggagcagcagcagcagcagcagcagcagcagcagcagcag 1022
 Qy 1021 cgtcagtggtctcaaaaggaaccatccttgaggccacgagcagcagcagcagcagcagc 1080
 Db 1023 cgtcagtggtctcaaaaggaaccatccttgaggccacgagcagcagcagcagcagcagc 1082
 Qy 1081 gccctglaagctggcagcgtaccctccgcccagcagcagcagcagcagcagcagcagc 1140
 Db 1083 gccctglaagctggcagcgtaccctccgcccagcagcagcagcagcagcagcagcagc 1142
 Qy 1141 actgctccggggccacagtcacatgctcctggtgagcagcagcagcagcagcagcagc 1200
 Db 1143 actgctccggggccacagtcacatgctcctggtgagcagcagcagcagcagcagcagc 1202
 Qy 1201 aggcacctacacccctgcctgtggaactccgctgctgctgctgctgctgctgctgct 1260
 Db 1203 aggcacctacacccctgcctgtggaactccgctgctgctgctgctgctgctgctgct 1262
 Qy 1261 gtagcgtggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1320
 Db 1263 gtagcgtggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1322
 Qy 1321 ctactcgtgctacagccgagcagcagcagcagcagcagcagcagcagcagcagcagc 1380
 Db 1323 ctactcgtgctacagccgagcagcagcagcagcagcagcagcagcagcagcagcagc 1382
 Qy 1381 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1440
 Db 1383 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1442
 Qy 1441 ccggcggcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1500
 Db 1443 ccggcggcggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1502
 Qy 1501 gtagcgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1560
 Db 1503 gtagcgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1562
 Qy 1561 gaataagactgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1620
 Db 1563 gaataagactgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1622
 Qy 1621 tgtggtctccaaaggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1680
 Db 1623 tgtggtctccaaaggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1682
 Qy 1681 ccccgagcgtctacacatcgaaatccaaagcagcagcagcagcagcagcagcagcagcag 1740
 Db 1683 ccccgagcgtctacacatcgaaatccaaagcagcagcagcagcagcagcagcagcagcag 1742
 Qy 1741 gctcctgagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800
 Db 1743 gctcctgagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1802
 Qy 1801 cctgtccacgctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1860

Db 1803 cctgttcacgctgcacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1862
 Qy 1861 tctgttcacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacg 1920
 Db 1863 tctgttcacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacg 1922
 Qy 1921 cagcgtcagcgtgagtcacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1980
 Db 1923 cagcgtcagcgtgagtcacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 1982
 Qy 1981 agtgcagacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacg 2040
 Db 1983 agtgcagacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacgacg 2042
 Qy 2041 cctggaagcctcctggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
 Db 2043 cctggaagcctcctggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2102
 Qy 2101 qctggaagcctcctggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2160
 Db 2103 qctggaagcctcctggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2162
 Qy 2161 cgaagagcgtcgtgaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2220
 Db 2163 cgaagagcgtcgtgaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2222
 Qy 2221 cctcagcgtcgtgaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2280
 Db 2223 cctcagcgtcgtgaggaagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2282
 Qy 2281 ggcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2340
 Db 2283 ggcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2342
 Qy 2341 ggcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2400
 Db 2343 ggcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2402
 Qy 2401 cctcagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2460
 Db 2403 cctcagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2462
 Qy 2461 cctcagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2520
 Db 2463 cctcagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2522
 Qy 2521 tgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2580
 Db 2523 tgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2582
 Qy 2581 cgcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2640
 Db 2583 cgcgtcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2642
 Qy 2641 cagcgtgagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2700
 Db 2643 cagcgtgagcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2702
 Qy 2701 gtcgagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2760
 Db 2703 gtcgagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2762
 Qy 2761 ggcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2820
 Db 2763 ggcgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2822
 Qy 2821 cctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2880
 Db 2823 cctcctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2882
 Qy 2881 cagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2940

[illegible]

Tue Jan 23 12:06:46 2001

Db 146 CACAGAGGATTATATGTGATTGACACCGGGGACAGACGCTATTCATATATCTTCAGGGGACA 205
 Qy 181 gacccccctcagtggtgttgccagagagctcagagagcgccagccaccgagagacaagga 240
 Db 206 GCACCCCTCTGAGTGCAGCTGGCCAGGGGCCAGGAGTACTGACACACAGGTGGGAAGGA 265
 Qy 241 cagcagagacacgggggtggtgagagactgcagagggcacagacgacgagccctactcaaa 300
 Db 266 CAGTGAAGACACACGGGTGTGATGATGTGAAGGCACAGAACTAGGCCCTACTGCAA 325
 Qy 301 ggtgtgtgctgcagagaggtacatgcacacgacagcagctacgtctgtactacaa 360
 Db 326 GGTGTACTGTCTGGCCAGACTCACGCCACACACACAGGGCAGCTACCACTGCTACTACAA 385
 Qy 361 gtacatacaggcagcagctcagagggccacacgcccgcgcagctcctacgtgtgtcgtgagaga 420
 Db 386 GTACATCAAGGCCGGATTTGAGGGCACCACAGCTGCCAGCACCTATGTGTTGTAAGAGA 445
 Qy 421 ctltgagagccatcacaacagcctgcacagcctcttggltcaacaggaagagccat 480
 Db 446 CTTTAAACACCCCTTTTCATCAACAACTGACACAGCTCTCGTGTCAACAGGAAGGACTCGAT 505
 Qy 481 gtgggtgcctgtctgtgtccatccccgcgcctcaatgtcaagctgcgctcgcgaagctc 540
 Db 506 GTGGGTGCCCTGTTGTTGCCATTTCCGGCCTCAACATCACACTGCGCTCGCAAAAGCTC 565
 Qy 541 ggtgtgtggtgcagagcagagaggtgtgtggtgagcagcgcggtgcgtctgtgtc 600
 Db 566 AGCCGTGCACCCCGACGGGAGGAGGTGCTGTGGGATGACCCGGGGCATCCGGGTGCC 625
 Qy 601 cagcgaactgtgcagatgcctgtacctgcagtcgagagacacactggggagccagga 660
 Db 626 CACTCAACTCTTGGCGATGCTGTACCTGTGAGTGCAGTCCGAGACACCTTGGGGTGACAGAA 685
 Qy 661 ctctcttcaaccccttctgtgtgacatcaacagcagcagcagctctatgacatccagct 720
 Db 686 CTTCTTTTCAACTCTCTGCTGTGCATATCACAGGCAATGAGCTCTATGACATCCAGCT 745
 Qy 721 gttgccagagagtcgtgtgagtcgtgtgagggagagagtcgtgtgacactgcacgt 780
 Db 746 GTACCCCAAGAACTCAATGGAGCTGTTGGTTGGAGAGAAAGCTGTTTGAACCTGTACAGT 805
 Qy 781 gtgggtcagtttaactcagtggttacccttgactgggactacccaggaagcagagcaga 840
 Db 806 GTGGGTGAGTTCGACTCAGGTGTACCTTCGACTGGGATTTCCAGGAAGCAGGCACA 865
 Qy 841 gcggggtlaagtgggtgcagcagcagcgtcccgagcagacccacacagaaactctccagcat 900
 Db 866 GCGGGCTAAGTGGGTACTTGNCGCGCTTCCAGCAGACGCCACACAGAACTCTCCAGAT 925
 Qy 901 cctgacctcaacaagctcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 960
 Db 926 CCTGACCATCCACAATGTACAGCCAGAATGACCTGGGGCCCTATGTGTGTGAGGCCAACAA 985
 Qy 961 cggatccagcagatttcggagagcagcagcagcagcagcagcagcagcagcagcagcag 1020
 Db 986 TGGGATTCAGCGGTTCCGGGAAACACAGAGGTCTATTGTGCACGAAAGCCCTTCATCAG 1045
 Qy 1021 cgtcagtggtcctcaagagacccatctctggagggccagcagcagcagcagcagcagcagc 1080
 Db 1046 TGTGAGTGGGTCAAGAGCCGTCTCTGGAGGCCACAGCCGGTGACGAGCTGTGTGAAGCT 1105
 Qy 1081 gcccgtgagctggtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1140
 Db 1106 ACCCGTGAAGCTGGGAGGCTTTATCCCCACCCAGGATTCATGTTACAGGACAGAAAGGC 1165
 Qy 1141 actgtccgggcccacagtcacatgccctgggtgtcctcaagagagtgacagagcagcagc 1200
 Db 1166 ACTGACTGGGGCCACAATCCCATGCTCTGTTGCTCAAGAGGTGACCGAGGCCAGCGC 1225
 Qy 1201 aggcacactcacccctgtggaactcgtgtggtgtggtgtggtgtggtgtggtgtggtgtg 1260
 Db 1226 AGGGGTCTACTCTCGCCCTGTGGAACTCTGCACTGTGAGGCAAAACATCAGTCT 1285

Qy 1261 gaagctgggtggat 1320
 Db 1286 GGAGCTGTGTGTGATGTGCTCTCATCATCCAAAGGAAGGCTCTTACCCAGCAT 1345
 Qy 1321 ctactcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1380
 Db 1346 CTACTCCGCCAGTGCAGACCCCTACCTGACCCGCTATGAGTACCCAAACCCCT 1405
 Qy 1381 cagcatccagtggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1440
 Db 1406 CAGTGTCCAGTGTGACACTGGAGGCCCTGGACACCCCTGCAAGACGTTTGGCCAGCGAGCT 1465
 Qy 1441 ccggtgcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1500
 Db 1466 CCGAGGGCGCAGCAGCGGATGGCATGCCAGTGCAGTCCGAGACTTGGAGAGGAGTCAACAC 1525
 Qy 1501 gcaaggtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1560
 Db 1526 TCAGGATGTGTGAACCCCATTCAGAGTCTGGACAGCTGGACGGAGTTTGTGGAGGGGAA 1585
 Qy 1561 gaataagactgtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1620
 Db 1586 GAATAACAGCTGAGCAAGCTGGTATCCAGATGCCAATGTGTGACGCTGTACAGGTG 1645
 Qy 1621 tgggtgtccaaaggtgggag 1680
 Db 1646 TGTGCTCTCAACAAGTGGGAGGATGAGAGACTCTACTTCTTATGTGTGACCACT 1705
 Qy 1681 cccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1740
 Db 1706 CCGGAGCGTTTCACTATCAGTTCGAGGAGCTTCTGAGGATCCCTTAGAAGGCCAGTCCGT 1765
 Qy 1741 gctcctgagctgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1800
 Db 1766 CGGCTCAGCTGCGGGCGGACAACTACAGTACGAAACATCTGCGCTGTGTGACCGGTCAA 1825
 Qy 1801 cctgtccagcgtgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1860
 Db 1826 CCTTCCACACTGCAGATGCTCAAGGAAACCCCTATTCTGTGAGTGCACAAACAGTCA 1885
 Qy 1861 tctgtccagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1920
 Db 1886 CCGTGTTCACAGCCCTTAGAGGCCAACCTAGAGGAGGAGAGAGCCCGGGGCCCGCCACGC 1945
 Qy 1921 cagcgtcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1980
 Db 1946 CACCTCAGTTGATATATCCCGGAGTGGCGCCGAGGAGGAGGAGTGTGACTAGTGTGTGA 2005
 Qy 1981 agtcaagacccggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2040
 Db 2006 AGTGCAGGATAGCGCAGCCAGGACAGCTTGGCCACAAAGATACCTTGTCCGTGTGAGGC 2065
 Qy 2041 cctggaagccctcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2100
 Db 2066 CCGTGAAGTCTCTCGGCTACCGCAACTTACCGACCTCTGTGTGACGCTGTGAGTGTGACTC 2125
 Qy 2101 gctgagagatgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2160
 Db 2126 CCTGGAGATGCGATCCCGGTGGTGGAGCGCATGTGCCAGTATTGTGTGTGTACAAAGA 2185
 Qy 2161 cagagagc 2220
 Db 2186 TCAAAAGCTCTTGGAGAAAGAGTGGGAAATCCACCTGGCAGACTCCAATCAGAGCTGAG 2245
 Qy 2221 catccagc 2280
 Db 2246 CATCCAGCGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2305
 Qy 2281 gggcgtcgtcaactcctccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2340
 Db 2306 GGGCTCGTAAACTCTCTGCCAGCGTGGCAGTGGAGGAGCTCTGAAGATAAAGGAGCAGCAT 2365

QY	2341	ggagatcgtgatacttgcgtacccgcgtacatcgctcttcttcttcttggtgctctccct	2400
Db	2366	GGAGATTGTGTAACATATTGCATCGCATCGCTATCGCAGTTTCTTCTTGCGTCCTCTCT	2425
QY	2401	cctcatcttctgaacaatgaggagcggccccacgcagacatcaagaagggtaacctgac	2460
Db	2426	GCTCATCTTCTGAACATGAAGAAGCCCTGCCATCGAGACATCAAGACGGCTACTCTGTC	2485
QY	2461	catcatatggacccccggggagggtgcctcttgaggagacaatgcgaataactctctaaga	2520
Db	2486	CATCATCATGGACC CGCGGAGGTGCCTTTTGGAGGACAGTGTGAATACTCTGCTATGA	2545
QY	2521	tgccagccagtgggaattccccagagacgcctgcacctgggagagtgctcgcgtacgg	2580
Db	2546	CGCCAGCCAGTGGGAGTTCCCCAGGAAGGTTGCACCTCGGAGAGTCTTAGGCACGG	2605
QY	2581	cgcttcgggaaaglggtggaagcctccgcttttcgcatcaccaaggcagcagctgtga	2640
Db	2606	GGCTTTTGGGAAGTGTGGAAGCCTCAGCTTTTGGCATCAATAAAGCAGCAGCTGTA	2665
QY	2641	cacgtggccgtgaaaatctgaaaaaggggccgacagcagcgagacacgcgcgctgat	2700
Db	2666	CACCGTGGCTGTGAAGATGCTGAAAGAGGGCGCTACTGCCAGCGACACCGTGCCCTGAT	2725
QY	2701	gtcgaagctcaagatcctcattcacatcgcgaacacacctaacgctggtcaacctctcgg	2760
Db	2726	GTCGGAGCTCAAGATCCTAATTTACATCGCACACCATCTCACGTGCTCAACTCTCTAGG	2785
QY	2761	ggcgtgcaccaaacgcgagggccccctcatggtgatcgtggagttctgcaagtaecggcaa	2820
Db	2786	GGCGTGCACCAAGCCCAACGGCCCTCTCA'TG'GTATCGTGGAGTTTGGCAAATACGGCAA	2845
QY	2821	cctctcaaatctctgcgcgcaagcggaacgccttcagccccctgcgcgagaagatctcc	2880
Db	2846	CCTCTCCAAGTCTTCGTGTCAACGGGACACGTTTTAACCCCCTACGCGGAGAAGTCTCC	2905
QY	2881	cgagcagcgcgacgtctcgcgcctatggtgagctgcacagcgctgga tcggaggcgcc	2940
Db	2906	GGACAACGCAGCGGCTTCGGCGCCATGGTAGAAGGCCCAAGCTGATAGGAGGAGACC	2965
QY	2941	ggggagcagcgcaggggtcctcttcgcgcggttctcgaagacgcgaggcgagcgaggcg	3000
Db	2966	TGGAAGCAGCGACAGGGCCCTGTTACCGCGGTTCCTGATGGCAAAGAAAGTGCACGGCG	3025
QY	3001	ggcttctccagacaagaagctlgaagacctlgaactaagcccgctgacctggaagatct	3060
Db	3026	AGCCCCACTTGTCCAGAGCTGAGACCTATFGCTGAGCCCCACTGACCATGGAAGACCT	3085
QY	3061	tgtctgtacagcttccaggtgccagaggatgagttcctggtctcccgaaagtgcatt	3120
Db	3086	TGTATGTACAGCTTCCAAGTTGCCCGGGAATGSAGTTCTGTGCTTCCCGCAAGTGCAT	3145
QY	3121	ccacagagacctggctgctcoggaaactctcgtctcgaaagcagcgtggtagaatctg	3180
Db	3146	TCACAGAGACCTGGCTCTCGGAACATCTCTGTCTGAAAAGTGCACATAGTGAAGATCTG	3205
QY	3181	tgaacttggccttgcgcgagacatclacaagaacctgaactlacgtlcgcgaaggcagctgc	3240
Db	3206	CGACTTTTGGCCCTGCTCGGAGCATCTACAAAAGCCCCGACTATATGTCCGAAGGGCAGTGC	3265
QY	3241	ccgcgtgccctgaagtggatggccctgaaagcatacttcgacaaggttacaccaagca	3300
Db	3266	CGGACTTCTCTGAAATGGATGGCCCCCGAGAGCATCTTTGATAAGGTGTACACACAGCA	3325
QY	3301	gagtgacgtggttccttggggjgctctctcggagagatcttctctcgtggggcctcccc	3360
Db	3326	GAGTGATGTGTGCTCTTCGGCGCTGCTGTGTGGAGATCTTCTCA'I'IGGGGGGCTCTCC	3385
QY	3361	tlacctctgggtgcagatcaatgagagttctcgcagcgcgctlgagagacgcacaaagat	3420
Db	3386	ATACCTCTGGGTACAGATCAATGAGAGTTCTGCCACGGCTGAAGGATGGCACTCGAAT	3445
QY	3421	gaggggccccgcgagctggccaactcccccaatacgcgcgcatcatgtgtaactgctggtccgg	3480

Db	3446	GAAGAGCCCGGAACATGGGCACATCTTCCCATACGCCACATATGCAGAGTTGCTGGTCTGG	3505
Qy	3481	aaaccccaagcgagacctgcaattcggagctcggtagatccctggggacacctgctcca	3540
Db	3506	AGACCTTAAAGCAAGACCTGCTTCTCTGACTAGTGAGATCTCTGGGACACCTGCTTCA	3565
Qy	3541	gggcaggggacctgcaagaggaagagaggttgcattgggccccgcgagctctcagaagctc	3600
Db	3566	GGCGGAGGCTTGGCAGGAGGAGGAAGAGGCGCATGGCCCTGCACAGCTCTCAGAGCTC	3625
Qy	3601	aaagagggcagctctctcgacaggtgtccacattggccctcacatcgcaccagcgtgcagc	3660
Db	3626	CGAGGAGGATGGCTTCATGACGGCATCCACCACAGCTCTACATATCACCGAAGCAGACGC	3685
Qy	3661	tgaggacagccgcgcaagcctgcagcgccacagcctggccgcccaggtatttacaactgggt	3720
Db	3686	TGATGATAGTCCACCCAGCATCATTTGGCACAGGCTGGCAGCCAGATATTACAACTCTGT	3745
Qy	3721	gtccttcccgggctcctggccagagggcttgagaccgctggttccctccagatgaagac	3780
Db	3746	GTCTTTCTGGCGGCTTGGCCAGAGGCNCTAAGACTCCAGGCTCTCCAGGATGAAGAC	3805
Qy	3781	atttgaggaaattccccatgaccccaagcactacaaggctctgtggacaaccagacaga	3840
Db	3806	ATTTGAAGAAATCCCATGACCCCTACAACCTTCAAAGGCTCCATGGATAACACAGACAGA	3865
Qy	3841	cagtggagtggtgtgctcctcgagagagtttgacagatagagagcagcgtatagacaaga	3900
Db	3866	CACGGGATGGTCTGGCTCAGAAGAGTTTCAGGAGCTAGAAAGCAGGCATAGACCAGA	3925
Qy	3901	aagcggttcagctgttaaaggacctggccagaaatgtggctgacccagggcacacctga	3960
Db	3926	AGCAGCTTCAGCTGTAAAGGTCCTGGCCAGGACACATGGATATTCACGAGGACACCTTGA	3985
Qy	3961	ctcccaaggagcgcgcgccctgagcggggggcccgagagggcagggtgttttcaaa	4020
Db	3986	CCCCAGGGGAGCGCGACGGCCCCACTCAAGGGGCACAAAGGAGCAAGGTGTTTATAA	4045
Qy	4021	cagcgaataggggagctgctggagaccaagcagagaggaacactgctcccgtctgcccg	4080
Db	4046	CAACGAGTATGGGAGGCTCTCCGAGGCATGTACAGAAGGTGACTGCTGCCGTCTGCTGG	4105
Qy	4081	cgtgactttcttcacagacaacagctactaa	4111
bh	4106	CTCCACCTCTCTCGACAGCAGCAGCTACTAA	4136

RESULT	8	
CCQUEK2		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
FEATURES		

3506	tcggagctggtgagatcctctggggagacgtctccaggcagggcctgcagaaggaagag	3565
3571	TCAGACCTGGTGGAGATACTTGGGAACCTCTCTCAGGAAATG---TCCAAACAGGAAGGG	3627
3566	gaggtctcatggcccccgcgcagctctcagagctcagaagagggcagctctctcgcaggtg	3625
3628	AAGGATTACATCCCACTGAATGACTCTCACAGCTCAGAAGATGATGGTTTCTCCACAGTG	3687
3626	tcacacatggcctcacatcgccagggctgacgctgagcgtgagcagacccgcgaacclgcay	3685
3688	CCTTCCTCTGCCCGACAAAATTCAGATGAAGAAGACTTTGCACATG-----AGGATACGT	3741
3686	cgccacagcctggccgcaggtattacaactgggtgctcttcccggtgctcgccgacaga	3745
3742	TGTCACAGCCTAGCAGCAAGATATTACAACCTCGCTCTCGTTCCCTGGTTGTTGCACGGGA	3801
3746	ggggctgagaccctggttctccaggatgaagacatttgaggaattccccatgacccca	3805
3802	GGAAATCAGATCAGGTGTTTCATCCAGGATAAAGACATTCGAAGAGTTTCCCATGACACAT	3861
3806	acgacctacaaaggtctctgtgacacacccagacagacagtggtgctggtgctcggag	3865
3862	ACAATGTACAAGGCACACCCGGACAATCAGACAGACAGTGGGATGGTTTGGCATCTGAG	3921
3866	gagttgagcagatagagcagcagtagacaagaagcggcttcagctgtgtaagagacct	3925
3922	GAATTTGAGAGGATAGAAACCGACACAGAAAGAGGTGGATTCAGCAGCAAGGGGCC	3981
3926	ggccagatgtggtctgtgaccagggcacacctgactcccaaggagcgcgcggcgcct	3985
3982	AACCGAACTCGGAGCTGTTCACAGCAACAGTCAGACCTCGCGGGCAGATGTCGGCCGTCA	4041
3986	gagcggggggccgaggaagccaggtgttttacaacagcagatgagggagctgtcggag	4045
4042	TACGGATCCAGTCCGAGGCCAGACTTTTACAACAGTGAATACGGGAGCTGTGCAGAA	4101
4046	ccaagcagaggaagaccactgctcccgctctgc	4077
4102	CACCTCTGAGGACCGCAGCTGCACCCACCCGCG	4133
RESULT 9		
LOCUS	144515/c	
DEFINITION	Sequence 17 from patent US 5635177.	PAT 07-OCT-1997
ACCESSION	144515	
VERSION	144515.1	GI:2469228
KEYWORDS	Unknown.	
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	1 (bases 1 to 6827)	
AUTHORS	Bennett, B.D., Goeddel, D. and Matthews, W.	
TITLE	Protein tyrosine kinase agonist antibodies	
JOURNAL	Patent: US 5635177-A 17 03-JUN-1997;	
FEATURES	Location/Qualifiers	
source	1..6827	
BASE COUNT	1558 a 1870 c 1738 g	1660 t 1 others
ORIGIN		
Query Match	25.3%;	Score 1040.2; DB 81; Length 6827;
Best Local Similarity	98.3%;	Pred. No. 3.1e-141;
Matches 1051; Conservative	0;	Mismatches 18; Indels 0; Gaps 0;
QY	2846	cggagcgtctcagccctcgccgagagaagctctcccgagcgcgcagcgtctccgcgc 2905
Db	2947	CGAGCCAGCTGTGTCATCCCGCAGGAGAAGTCTCCGAGCAGCGGACGCTTCCCGGCC 2888
QY	2906	atggtgagctcgcagcgtggtatcgagcgcgcggggagcagcagcagcagcagcgtcttc 2965

Db	2887	ATGTTGAGCTCCGCCAGGCTGGATTCGAGCGCGCGGGAGGACGACAGGGTCTCTTC	2828
QY	2966	qgcgggttctcgaagaccgagggcgagcgagcgagcggttctccagaccaagaagctgag	3025
Db	2827	CGCGGTTCTCGAAGACCGAGGCGGAGCGGAGCGGGCTTCTCCAGACCAAGAAGCTGAG	2768
QY	3026	gacctgtggtgagcccgctgaccatggaagatcttgtctgtctacagcttccaggtggcc	3085
Db	2767	GACCTGTGGTGTGAGCCCGCTTGACCATGGAAAGATCTTGTCTGCTTACAGCTTCAGAGTGGCC	2708
QY	3086	agagggatggagtlccttggcttccccgaagtgcattccacagagacctggctgctcggaac	3145
Db	2707	AGAGGGATGGAGTTCCTTGGCTTCCCGAAAGTGCATCCACAGAGACCTTGGCTGCTCGGAAC	2648
QY	3146	attctgtctcggaagcgcgcgtggtgaagatctgtgaactttgaccttgcscggagacatc	3205
Db	2647	ATTCTGTCTGCGAAAGCGACGTGGTGAAGATCTGTGACTTTGGCCTTGCCTCGGGACATC	2588
QY	3206	tacaaagacctgactacgtcccgcaaggcagtgacctcccggtgcccctgaaagtggatggcc	3265
Db	2587	TACAAAGACCCGACTAGTCCCGCAGGGCAGTGCCTGGGTGCCCTCTGAAGTGGATGGCC	2528
QY	3266	cctgaagacatcttcgacaaggtatacacacgcagaagtaacgtgtgcttcttgggggtg	3325
Db	2527	CCTGAAGACATCTTCGACCAAGTGTACACCACGACGAGTGACGTGTGCTCTTTGGGGTG	2468
QY	3326	cttctctgggagatcttctcttggggcctccccgtacctcccggtgggtggagatcaatgag	3385
Db	2467	CTTCTCTGGGAGATCTTCTCTGGGGCCTCCCGGTACCTTGGGGTGACAGTCAATGAG	2408
QY	3386	gagttctccagcggctgagagacggcacaaggatgagggcccgagctggccactccc	3445
Db	2407	GAGTTCTCCAGCGGCTGAGAGACGGCACAAAGATGAGGGCCCGGAGCTGGCCACTGCC	2348
QY	3446	gccatacgcgcacatcgtgaactgctlggtccggagagaccccaagcagagacctgcattc	3505
Db	2347	GCCATACGCCGCATCATGCTGAATGCTGGTCCGGAGACCCCAAGCGGAGACCTTGCATTC	2288
QY	3506	tcggagctgggtgagatccttggggacacctgctccaggcagggcgctgcaagaggaagag	3565
Db	2287	TCGGAGCTGGTGGAGATCCTGGGGACCTGCTTCCAGGCGAGGGCCTTGCAGAGGAAGAG	2228
QY	3566	gaggtctgcagtgcgccccgcagctctcagagctcagaagagggcagcttctcgcaggtg	3625
Db	2227	GAGGTCTGCATGGCCCGCCGAGCTCTCAGAGCTCAGAAGAGGGCAGCTTCTCGCAGGTG	2168
QY	3626	tccacatggccctacacatcgcccgagcgtgacgctgagagcagcccgccaaagcctgcag	3685
Db	2167	TCCACCATGGCCCTACATCCTCCAGGCTGACGCTGAGGACAGCCCGCAAGCCTGCAG	2108
QY	3686	cgccacagcctggcgccaggtattacaaactgggtgctcttccccgggtgctcgggccaga	3745
Db	2107	CGCCACAGCTGGCCGCCAGGTATTACAACTGGGTGCTCTTCCCGGGTGGCTTGGCCAGA	2048
QY	3746	ggggctgagaccgtggttcttccaggatgaagacatttgaggaattccccaatgacccca	3805
Db	2047	GGGGCTTGAGACCCCGTGGTTCCTCCAGGATGAAGACATTTGAGGAATTCCTCCATGACCCCA	1988
QY	3806	acgacctacaaaggtctctgtggacaccagacagacagtggaatggtgtgctggcctcgag	3865
Db	1987	ACGACCTACAAAGGCTCTCTGGACAACACAGACAGAGTGGATGCTGTGCTGGCCTCGGAG	1928
QY	3866	gagtttgaacagatagagagcagcagcatagacaagaacggttcaagct	3914
Db	1927	GAGTGTGAGCAGATAGAGCAGGTATAGACAAAGAAAGCGGCTTTCAGGT	1879
RESULT 10			
MMFLK1M	LOCUS	MMFLK1M	5470 bp
DEFINITION	M.musculus Flk-1 mRNA.		
ACCESSION	X70842		
VERSION	X70842.1		
	GI:57923		
		02-SEP-1996	

endothelial cell growth factor protein; Flk-1 gene; tyrosine kinase receptor.
Mus musculus
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
CDS
BASE COUNT
ORIGIN
Query Match 17.4%; Score 716.2; DB 11; Length 5470;
Best Local Similarity 54.6%; Pred. No. 1.7e-94;
Matches 1713; Conservative 0; Mismatches 1343; Indels 81; Gaps 11;
Qy 464 aacaggagagccatggtggtgcccgtctggtgtccatccccgcctcaatgtcacg 523

712 AACAAACAACAAACATG1GG1GATCCCTGCCGAGGGTCGATTTCAAACCTCAATGTGCT 771
Qy 524 ctgcgc-----tcqcaaaagctcgtgctgctgagccagagcagagaggtggtggtg 577
Db 772 CTTTGCGCTAGTATACCAAGAAAGAGATTTGTTCCGGATGGAACAGATTTCTCTGTC 831
Qy 578 gaccgaggggcatgctgctgctccacgcactgctgcagcagtgccctgactcagtcg 637
Db 832 AGGAGATAGCTTTTAC1CTCCCAAGTTACATGATCAGTATGTCGCGCATGCTCTCT 891
Qy 638 gagcacacctggggagaccagagactcttccaacaccttctctggtgacacacagc 697
Db 892 GAGGCAAGATCAATGATGAAACCTATCAGTCTATCATGTATAGTGTGTTGTAGA 951
Qy 698 aacgagctctatgacacccagctgttgcacaggaagtcgctgagctgctgtagggag 757
Db 952 TATAGATTTATGATGTGATTC1GAGCCCCCGCATGAAATGAGCTATCTCCCGAGAA 1011
Qy 758 aagctggtcctgaactgcacccgctggtggtgagtttaactcaggtgtcacccttgactg 817
Db 1012 AAACCTTGCTTAAAT1GTACAGCGAGAACAGAGCTCAATGTGGGGCTTGATTTACCTGG 1071
Qy 818 gactaccacgggaagcagcgc-----agagcggggttaagtgggtgcccagcagcgc 869
Db 1072 CACTCTCCACCTTCAAAGTCTCATCATAAGAGATTTGTAACCCGGATGTGAACCTTT 1131
Qy 870 cccagcagacccacacacactctcc--agcatcctgacctccacacagctcagcagcac 928
Db 1132 CCTGGGACTGTGGGGAAGATGTTT1TGAGCACCCTTGACAATGAAAGTGTGACCAAGAT 1191
Qy 929 gacctgggtcgtatgctgcaagggcccaacagcgcatccagcgatttcggggagagcacc 988
Db 1192 GACCAAGGGGAATACACCTGTGTAGCCTCCAGTGGCGGATGATCAAGAGAAATAGAAC 1251
Qy 989 gaggtcattgtgatgaaatcccttcacatcagcgtgagtggtcctcaagagaccatctg 1048
Db 1252 TTGTCCGAGTTACACAAAGCCTTT1ATTGCTTCGTAGTGGATGAAATCTTTGGTG 1311
Qy 1049 gagccacggcagagcagcagctggtgaaagctgcccctgaagctgagcagctaccctccg 1108
Db 1312 GAAGCCA---CAGTGGGCGAGTCAAGTCCGAATCCCTGTGAAGTATCTCAGTTACCCAGCT 1368
Qy 1109 cccagttccagtggtacaaggtgaaagcactgtccggcgccacca-----gt 1159
Db 1369 CCTGATATCAATGGTACAGAAATGGAAGCCCATTTGAGTCCAACTACACATGATTGT 1428
Qy 1160 ccacatgccctggtgctcaagagggtgacagagcagcagcagcaccacacacctgcgc 1219
Db 1429 GCGCATGAATCACCATCATGCAAGTAC1GAAAGAGATGCAAGAAATACACGGTCAATC 1488
Qy 1220 ctgaggaactcgcgtgc1ggcctgagcgcaacatcagcctggagctggtgagtgatg 1279
Db 1489 CTCACCAACCCCATTTCAATGGAGAAACAGACCCACATCGGTCCTCTCTGTTGTAATGTC 1548
Qy 1280 cccccccagatatacatgagagagcctcctccccc-----geatctactcgcgcacagc 1336
Db 1549 CCACCCACAGATCGGTGAGAAAGCCTT1GATCTCGCCTATGGATTCCTACCATGATGGGACC 1608
Qy 1337 cgcagggcctcaactgcacggcctacaggggtgcctcctcctcagcatcagtgagc 1396
Db 1609 ATGCACACATGACATGACAGTGCACAGCTACGCCCAACCCCTCCCTGCACACATCCAGTGGTAC 1668
Qy 1397 tggcgccctggacacccctgcaagatg1ttgcccagcgtagttctccgcgcgcgcgcag 1456
Db 1669 TGCACCTACAAGAACGCTTC-----TCCTACACACCCGGCCAAA 1708
Qy 1457 caagacctcaatccacagtcgctgactgagggggcggtgacctgacacgcagatcccggaac 1516
Db 1709 CAAGCCCGTATGC---TTGTAAGAAATGGAGACAGCTGGAGGATTTCCAGGGGGGAAC 1764
Qy 1517 cccatcgagagcctggacacacctgaccgagatttggagggaagaataagactgtgacg 1576

Db 1765 AAGATCGAAGTACCAAAACCAATATGCCCTGATTGAAGGAAAAACAAACAACTGTAAGT 1824
Qy 1577 aagctggatgacgaatgcaaaagtgctccatgtacaaagtgatgagtgctcacaacag 1636
Db 1825 AGCTGGTTCATCAAGCTGCCAACGTGTACAGGTGTGAAGTGTGAAGCCATCAACAAA 1884
Qy 1637 gtggccagggatgagggctcactactctctctctctctctctctctctctctctcacc 1696
Db 1885 GCGGACGAGGAGAGAGGGTCTCTCTCCATGCTGATGATCAGGGTCTCTGA-----A 1935
Qy 1697 atcgaatccaagccatccagagagactactadagggccagccggtgctcctgaagtgccaa 1756
Db 1936 ATTACTGTCAACCTGTCTGCCAGCCAACTGAGCAGGAGATGTGTCTCTGTGTGCACT 1995
Qy 1757 gccgacagctacaagtaagagcatctgctgctggtaccgctcacaactgtccacgctgac 1816
Db 1996 GCAGACAGAAATACGTTTGAGAACCTCACGTGTGTACAAAGCTTGGCTCACAGCAACATCG 2055
Qy 1817 gatgcgacgggaacccgcttctgctgactgcaagaagctgcatctgttcgcacccct 1876
Db 2056 GTCCACATGGGGCAATCACTCACACAGTTTGCAAGAACTTGGATGCTCTTTGGAAACTG 2115
Qy 1877 ctggcccgagcgtgagaggtggcactggtggtggtggtggtggtggtggtggtggtggt 1936
Db 2116 AATGGCACCA-----TGTTTCTAACACACAAATGACATCTTGATGTGGCA 2163
Qy 1937 atccccgcgctgcgcccagacgagggccactatgltgcgaagtgcgaagaccggcgc 1996
Db 2164 TTTCAAGATGCTCTCTGCGACCAAGCGACTATGTTGCTCTGCTCAAGATGAAGAG 2223
Qy 1997 agccatgacaagctgcacaaagctgctgctgctgctgctgctgctgctgctgctgctgct 2056
Db 2224 ACCAGAAAGACATGCTGTGTCAACAGCTCATCTCTAGACGGCTGACACCCATG 2283
Qy 2057 ctccagcagaacttgacgacactcctggtgaacgtgagcactgctgagatgcagtgc 2116
Db 2284 ATCACCAGAAATCTGGAGAACTCAGACAAACCAACCAATTTGGAGACCACTTGAAGTGA 2343
Qy 2117 tgggtggccgagcagc 2176
Db 2344 CCAGCATCTGGAATCTTACCCCAACATATACATGTTTCAAGAACAGACGACCTTGTA 2403
Qy 2177 gaaagcttgagctgagctgagctgagctgagctgagctgagctgagctgagctgagctgag 2236
Db 2404 GAAGATTACAGCATTTGACTGAGAGATGGAAACCGGACCTGACTATCCGAGGTGAGG 2463
Qy 2237 gaggagatgcggagcgtatctgtgagcgtgtgcaagcgcgcgcgcgcgcgcgcgcgcgcgc 2296
Db 2464 AAGGAGGATGGAGGCTCTTACACCTGCCAGGCTGCAATGTCTTGGCTGTGCAAGAGCG 2523
Qy 2297 tccgcagcgtggcgtggaagctccagagataagggcagcagcagcagcagcagcagcagc 2356
Db 2524 GAGACGCTCTTCAATATAGAGGTGCCAGGAAACCTGAACTTGAAGTCAATATCTCTC 2583
Qy 2357 gtccgtaccggtcagcgtgct 2416
Db 2584 GTCCGCACTGCAGTATTGCCATGTTCTTCTGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 2643
Qy 2417 atgaggagggcccaagcagacatcaagacggcgtacctgtccatcalcatgagccccc 2476
Db 2644 GTTAAGCGGGCAATGAAGGGAACTGAAAGACAGGCTACTTCTTCTTCTTCTTCTTCTTCTTCT 2703
Qy 2477 ggggaggtgctctgagaggaatgcgaatccctgctcactgactgactgactgactgactgact 2536
Db 2704 GATGAATTCCTTGGATGAGCGCTGTGAACGCTTGCTTATGATGCCAGCAAGTGGAA 2763
Qy 2537 tcccccgagcgtgcacctgggagagtgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2596
Db 2764 TTTCCCGAGGACCGGTGAACTAGGAAACCTCTTGGCCGGGTGCTCTGCGCCAAAGTG 2823
Qy 2597 gtggaagcctccgtcttcgacatcacaaagggcagcagctgagacacgtgagccgtgaaa 2656
Db 2824 ATTGAGGCAGACCTTTTGAAATTGCAAGACAGGACCTTGGGAAACCTCTCTGCAACAAATCGG 2883

Qy 2657 atgctgaagagggcgccacggccagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2716
Db 2884 ATGTTGAAAGAGGAGCAACACACAGCGAGCATCGAGCCCTCATGTCTGAACTCAAGATC 2943
Qy 2717 ctaattcacatcgcgacccacctcaacgtgtaacctcctcggggcggtgacacaaagcg 2776
Db 2944 CTCATCCACATTTGGTGAACCTCTCAATTTGGTGAACCTCTCTAGGCGCTTCCGCAAGCG 3003
Qy 2777 caggcccccctcatggtgactgaggtgctgcaagtlacggcaacacctctcaactctctg 2836
Db 3004 GGAGGCTCTCATGTTGATTTGCAATTTGCAAGTTTGGAAACCTATCAACTTACTTA 3063
Qy 2837 cgcgcaagcgggagccttcagccccctgcgagagtgctccccgagcagcggcgcgcgcgcgc 2896
Db 3064 CGGGCAAGAGAAATGAAATTTGTTCCCTATAAGAGCAAAAGGGGACGCTTCCGCGAGGC 3123
Qy 2897 ttcgcgcctatggtgagcgtccagcagcgtggtggtggtggtggtggtggtggtggtggtggt 2956
Db 3124 AAGGACTACGTTGGGAGCTCTCCGTGATCTGAAAGACGCTTGGACAGCATCACCAGC 3183
Qy 2957 gtccctcttcgcgcgttctcgaagaccgagggcgagcggggtctctccagaccaa 3016
Db 3184 AGCCAGAGCTCTGCCAGCTCAGGCTTTGTTGAGGAGAAATCCCTCAGTGTAGAGCAA 3243
Qy 3017 gaag-----ctgaggaacctgtgctgagccccctgacctggaagatcttctgtctgtac 3070
Db 3244 GAAGAAGCTTCTGAAGAACTTACAAGGACTTCTCTGACCTTGGAGCATCTCATCTGTGTAC 3303
Qy 3071 agcttccaggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 3130
Db 3304 AGCTTCAAGTGGCTAAGGCTAGGCTTCTTGGCATCAAGAAAGTGTATCCACAGGAC 3363
Qy 3131 ctgctgctcgcggaacatctgctgctggaagcagcagctggtggaagatctgacttggc 3190
Db 3364 CTGCGAGCAGCAAACTCTCTATCGGAGAGAAATGTTGTTAAGATCTGTACTTCCGC 3423
Qy 3191 ctgccccgggacatctacaagacccctgactacgtccgcaagggcagtgccgcgctgcc 3250
Db 3424 TTGCCCCGGACATTTATAAGACCCCGGATTTATGTCAGAAAGGAGATGCCGACTCCT 3483
Qy 3251 ctgaagtggatggccccctgaaagcattcttcgacaaaggtgtacacacacagtgagctg 3310
Db 3484 TTGAAGTGCATGGCCCGCAAACTTTTCACAGAGTATACACAATTCAGAGCGATGTG 3543
Qy 3311 tggctcttgggggtgcttctctggtgagatcttctctggtggggcctccctgacctggg 3370
Db 3544 TGGTCTTTTCGGTGTGTCTCTGCGAAATATTTTCTTGGTGGCTTCCCTATACCTGGG 3603
Qy 3371 gtgcagatcaatgagaggttctgctccagcgtgagagcgcgcacagagatgagggcccg 3430
Db 3604 GTCAAGATTGATGAAGAAATTTGTAGGAGATTGAAGAGGAACTAGAAATCGGGCTCT 3663
Qy 3431 gagctggccactcccgccatacgc 3490
Db 3664 CACTACACTACCCAGAAATGTACCAGACCATGCTGTGACTGCTGGCATGAGACCCCAAC 3723
Qy 3491 gcgagacctgcatctcgcgagcgtgagatccttggggacccctgctccagggcagggc 3550
Db 3724 CAGAGCCCTCTGTTTCAGAGTTGGTGGAGCAATTTGGGAAACCTCTCTGCAACAAATCGG 3783
Qy 3551 ctgcaagaggaagagga 3567
Db 3784 CAGCAGGATGGCAAGA 3800

RESULT 11
AR068047
LOCUS AR068047 5470 bp DNA
DEFINITION Sequence 1 from patent US 5851999.
ACCESSION AR068047
VERSION AR068047.1 GI:5999269
KEYWORDS

Tue Jan 23 12:06:46 2001

[illegible]